POWER BI CHALLENGE

CRIME CASE

Group 10

Shanjida Roman – 20221395
Jaime Duarte – 20220675
Yousef Ebrahimi - 20221382
Mariana Lavinha – 20220565

TABLE OF CONTENTS

TABLE OF CONTENTS 1
1. INTRODUCTION 2
2. BUSINESS PROBLEM 2
3. BUSINESS QUESTIONS 2
4. DATA SOURCES 3
4.1. ENTITIES AND ATTRIBUTES 3
5. DIMENSION MODEL 4
5.1. RELATIONSHIPS BETWEEN ENTITIES 5
5.2. FACT TABLES 5
6. DISCOVERY PROCESS 6
6.1. EXTERNAL DATA AND TABLES 6
7. DATA MODELLING METHODOLOGY 6
8. DATA ENGINEERING 7
8.1. DESCRIPTION OF DATA ENGINEERING STEPS 7
9. MODEL OPTIMIZATION STEPS 10
10. DESCRIPTION OF MEASURES AND CALCULATED COLUMNS 11
10.1. DAX MEASURES 11
10.2. DATA FORMATTING: 14
11. ANALYSIS AND DISCUSSION 14
11.1. THREE MAIN KPIS 14
12. VISUALIZATION & MEETING PROJECT SCOPE 16
12.1. OVERVIEW 16
12.2. TIME INTELLIGENCE (TREND ANALYSIS 2018-2022) 17
12.3. GEOSPATIAL ANALYSIS 18
12.4. BY TYPE 19
12.5. ADVANCED VISUALS WITH AI FEATURE 20
13. ANALYSIS OF PROJECT OUTPUTS 21
14. CONCLUSION 22
15. RESOURCES 23
16. APPENDIX (LIST OF FIGURES AND TABLES) 24

1. INTRODUCTION

The Chicago Police Department (CPD) is one of the oldest and largest law enforcement agencies in the US. With over 13,000 officers and civilian employees, CPD is committed to maintaining law and order in Chicago. The department values integrity, professionalism, and respect, and works towards public safety and community partnerships. CPD's mission is effective policing, crime prevention, and creating a safe environment, while its vision involves utilizing innovative technology and tactics to reduce crime and increase public confidence.

2. BUSINESS PROBLEM

- How can we create a complete model that allows for the visualization and sorting of crime occurrences by type, district, and beat, and incorporates markers like arrests and domestic crimes to better understand Chicago's escalating crime rate?
- How can we evaluate past trends and patterns in crime data for months and years while making sure the model is flexible and dynamic enough to take into account fresh data and be used in future analyzes of other cities and states?
- What capabilities should be added to the PowerBI model to do this? How can we guarantee data privacy and confidentiality for users from various districts while providing them with access to pertinent crime statistics for their region?

3. BUSINESS QUESTIONS

1. Total number of crime events per day and event type sorting: The model should show the overall number of crime events per day and allow for event type sorting. The Chicago Police Department's Bureau of Records will be able to determine which types of crime events are more frequent and track their frequency over time as a result.

2. Understanding the scenario based on extra indicators: To better comprehend the circumstance, the solution must take into account additional indicators such as if someone was detained or whether the crime was committed within the home. This will assist the Chicago Police Department's Bureau of Records in identifying patterns and trends in crime and enhancing their approaches to combating it.

3. Determining the most affected districts and beats: To determine the most affected locations by crime occurrences, the model must assess the district and beat linked to each event. This will make it possible for the Chicago Police Department's Bureau of Records to use its resources more wisely and concentrate its efforts where they are most needed.

4. Using a map to depict crime hotspots and providing a table with names or labels for each district is part of the concept. This will make it possible for the Bureau of Records of the Chicago Police Department to recognize high-crime areas promptly and take the necessary measures.

5. The average interval between criminal acts: The answer must establish the typical interval between criminal acts and have the flexibility to take into account various kinds of criminal acts in this analysis. This will make it easier for the Bureau of Records of the Chicago Police Department to comprehend the trends and frequency of criminal events and allocate resources accordingly.
6. Recognizing changes over time: The model ought to compare crime data for months and years to examine prevailing trajectories and patterns. This will make it easier for the Bureau of Records of the Chicago Police Department to determine if a general crime is rising or falling and to comprehend the underlying factors.
7. Scalability for future growth: While initially focusing on Chicago, the solution should be flexible enough to take into account new data and apply the same analysis to other cities and states in the future. This will make it possible for the Chicago Police Department's Bureau of Records to increase its efforts to combat crime and communicate its top techniques to other law enforcement organizations.
8. Restricted visibility for users: To protect data privacy and confidentiality, the model should only allow users from one district to see information specific to that district. This will make it easier for the Chicago Police Department's Bureau of Records to adhere to data protection laws and make sure that private information is not disclosed to unauthorized parties.

4. DATA SOURCES

For this project, we have 5 data files, from 2018 to 2022. Each file provided by the Chicago Police Department (CPD) represents the crimes that occurred in Chicago in that year. We have 4 datasets including primary crimes type, wards representation, each Police Beat, and each Police District.

4.1. ENTITIES AND ATTRIBUTES

In this project, we have different entities and attributes. Entities: In the analyzing process we created attributes to answer the Police's questions. These entities are Police Beat, Police Districts, Wards, Crime Location, Primary Type, Crime data, and Shift Tables.

- Primary Type: Table which identifies the different types of primary crimes;
- PoliceDistricts: Table constituted by data about each Police District;
- PoliceBeat: Table composed of data about each Police Beat;
- Ward: Table that contains data relative to the wards' representation;
- Crime Data: A table that allows a more crime-type-oriented analysis and relevant details for each crime;
- Location Data: Here it is possible to analyze the crime from a geographical

perspective, having the possibility to indicate the exact coordinates of the crime location;

• Fact Crime Data: This table is only for a better organization of the content. It has all the foreign keys and measures.

5. DIMENSION MODEL

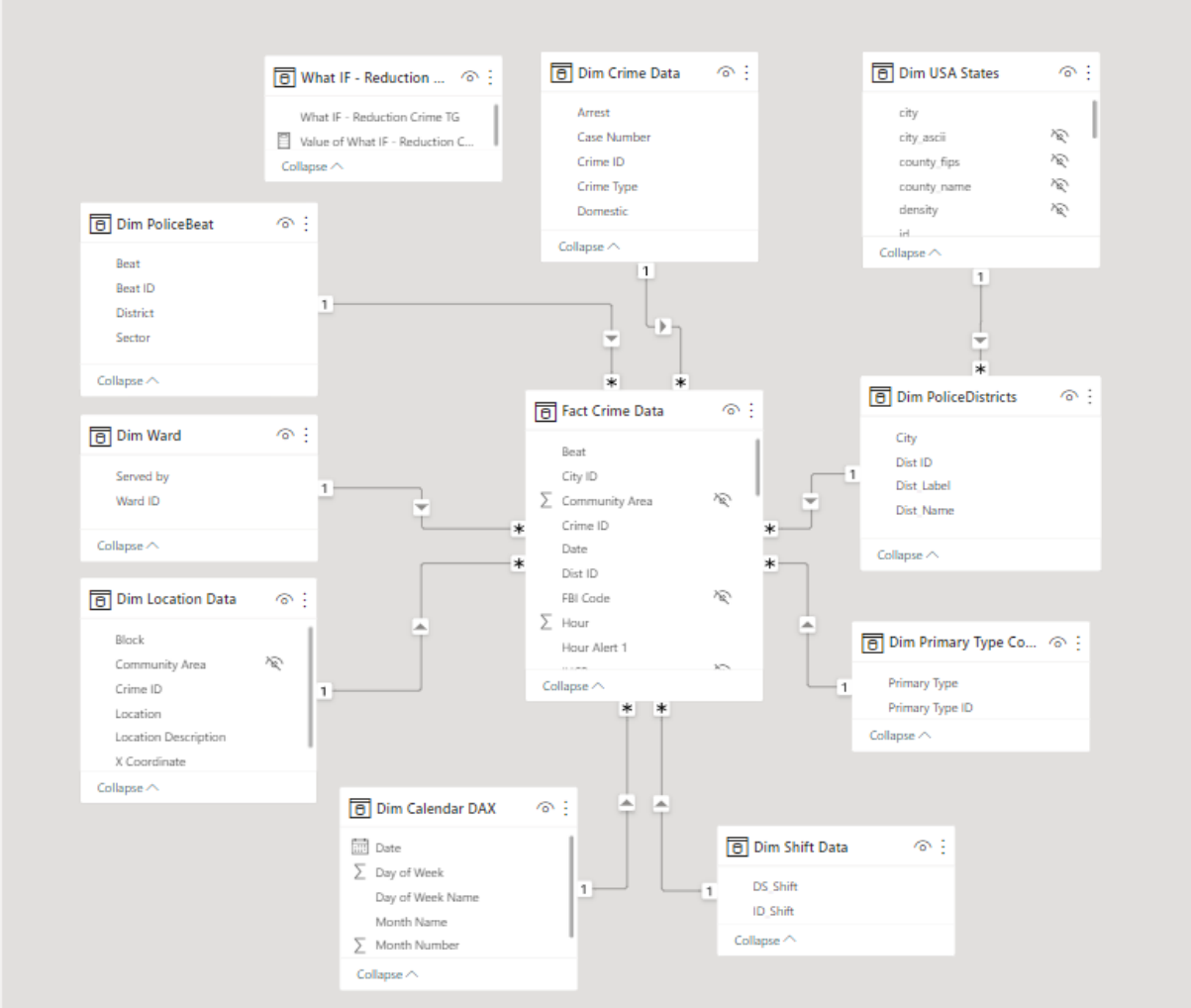


Figure 1: This figure shows the relationships of tables in our PowerBI file

In our model, we have identified nine dimension tables that are used to provide context to our data. These dimensions are Police Beat, Police Districts, Wards, Location, Primary Type, Calendar, Crime data, and USA States. Dimensions contain a hierarchy of attributes that enable users to drill down into the data and explore it at different levels of granularity. For example, the Date dimension includes attributes like year, quarter, month, and day, which allow users to analyze crime trends over different periods.

5.1. RELATIONSHIPS BETWEEN ENTITIES

All the entities are connected to the fact table with the relationship of one to many. For example, one record on the fact table can occur in only one city or location.

5.2. FACT TABLES

Our fact table is designed to capture the key metrics related to crimes in Chicago. It is connected to each dimension table using an appropriate relationship. The fact table includes measures such as the number of reported crimes, the number of arrests made, and the number of domestic crimes. Each measure is defined with a clear description of what it represents and how it is calculated. This ensures that users can easily understand the data and use it to derive meaningful insights.

Fact table also includes other various measures that are very important to calculate KPIs, cards, and graphs in reports and dashboards.

In summary, our model is based on the star schema, with nine dimension tables and a fact table that contains key metrics related to crimes in Chicago. This approach provides users with a flexible and intuitive way to explore the data and gain insights into crime trends in the city.

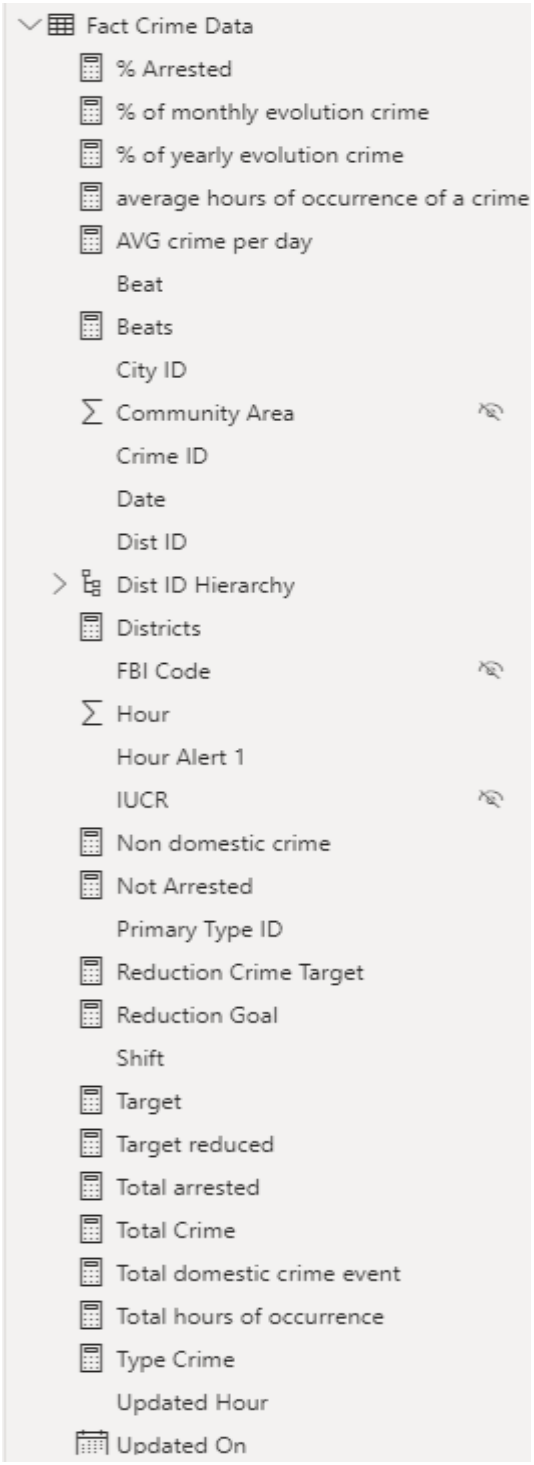


Figure 2: This figure shows the fact table properties.

6. DISCOVERY PROCESS

In our discovery process, we delved into various datasets, exploring details on primary crime types, police districts, beats, wards, crime locations, crime data, and dates. Our goal was to uncover valuable insights into the city's rising crime rates. By combining and analyzing this diverse information, we built a comprehensive model within PowerBI. This model allows us to visualize, sort, and analyze crime occurrences, giving us a dynamic and flexible solution. We didn't stop there. We made sure our model could incorporate fresh data and be used for future analysis in other cities and states. We aimed to equip the Chicago Police Department's Bureau of Records with a powerful tool that enhances their understanding of crime, optimizes resource allocation, and helps them devise effective strategies to combat criminal activities.

6.1. EXTERNAL DATA AND TABLES

During the discovery process, we realized that according to project requirements, we need some external data. Hence, two tables were created.

a. USA cities and population data were collected from the World Population Review website: worldpopulationreview.com
b. A new table **Dim Shift Data** was created to categorize police shift times.

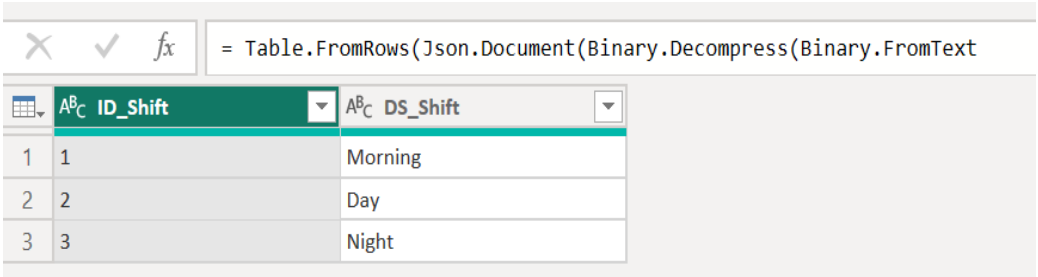


Figure 3: The table shows the police shift categories.

7. DATA MODELLING METHODOLOGY

We employed the renowned Kimball methodology, comprising four steps, to develop our data model for the Chicago Police Department (CPD). Initially, we focused on understanding the critical business problems and aimed to analyze crime patterns efficiently while optimizing resource allocation. We carefully determined the appropriate level of granularity for data collection and storage to ensure alignment with CPD's requirements. In the Dimensions step, we identified key perspectives, such as incident type, location, time, and police district, creating seven Dimension tables for comprehensive analysis. Lastly, we identified essential indicators, including occurrences, arrests, and domestic crimes, during the Facts step. The Kimball methodology provided a strong foundation for our Dimensional Model, enabling faster and more accurate crime data analysis for informed decision-making by the CPD.

8. DATA ENGINEERING

Data Engineering is a crucial aspect of the Crime project, and it involves the essential steps taken to prepare and transform the raw crime data for effective analysis and visualization in Power BI. We first did the cleaning the data to remove any inconsistencies or errors, integrated data from different sources to create a unified dataset, transformed the data to match the required format (like correcting the date format) which we will describe in the following, and loaded the data into the Power BI model.

8.1. DESCRIPTION OF DATA ENGINEERING STEPS

These steps involved refining and preparing various data tables to enhance their usability and improve the quality of the data. Each step is described, highlighting the specific actions taken to address data inconsistencies, standardize formats, and add new columns for analysis purposes.

Table 1: Describes the Data Engineering Steps by each table

Table	Description
Dim Primary Type	- Cleaned data by removing empty rows. - Transposed the table to convert rows into columns. - Filter null and white space - Capitalized each word
Dim Ward	- Enhanced data quality by converting column types to text. - Removed duplicate entries. - Split text using a delimiter. - Removed "Ward" from data and renamed columns for clarity. - Replaced null and 0's with a new ID that stands out for null Served by value.
Dim USA States	- Reordered states based on a specific criterion. - Adjusted column types accordingly.
Dim PoliceDistricts	- Renamed columns for clarity. - Ensured capitalization consistency.
Dim PoliceBeat	- Converted columns to the Int64 data type. - Replaced null values with 0 in the Sector column. - Checked and removed duplicated rows. - Removed redundant rows based on specific column values.
Illinois_Chicago_2018	- Transformed column data types into appropriate formats. - Corrected date format using locale for consistency and accurate data-based analysis.
Illinois_Chicago_2019	- Followed similar steps as Illinois_Chicago_2018.
Illinois_Chicago_2020	- Followed similar steps as Illinois_Chicago_2018.
Illinois_Chicago_2021	- Followed similar steps as Illinois_Chicago_2018.
Illinois_Chicago_2022	- Followed similar steps as Illinois_Chicago_2018.

Dim Crime Data	- Table created by appending yearly crime data file provided by the client - Filled null values in the Arrest and Domestic columns as instructed by clients on their info notes. - Removed some columns to streamline the data. - Fixed column names for clarity. - Removed duplicates.
Dim Location Data	- Table created by appending yearly crime data file provided by the client. - Deleted unnecessary columns for improved data clarity. - Ensured consistency and readability in the Location Description column by capitalizing words. - Removed brackets from the location column so that Mapping visual recognizes the location coordinates. - Replaced instances of null and "0" with "Unknown" in the Location Description column.
Fact Crime Data	- Cleaned data by removing leading or trailing spaces in cells. - Removed unwanted character (_) from Dist_ID column. - Added a new column "City ID" to each row. - Fixed column names for clarity. - Hide unused columns for this project that may be necessary for future use such as FBI code. - Added a condition column for shift data - Added a new column for the hour

For this project Date and time plays a crucial role. We took extra care in handling date variables.

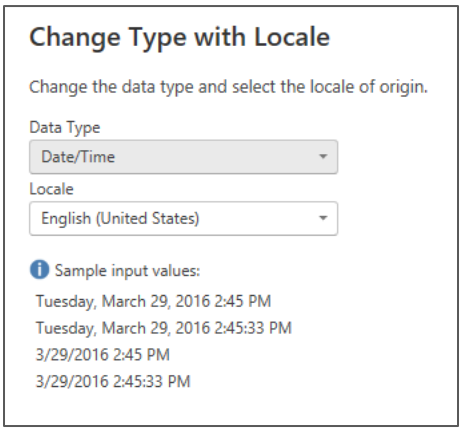


Figure 4: the date format used for all dimensions.

More details of these variables are below:

1. DIM Ward: null values were replaced by the code below:

```
#"Replaced Value2" = Table.ReplaceValue(#"Removed Duplicates”,
null,"9999",Replacer.ReplaceValue,{"Ward"}),
```

2. DIM Crime Data: null values were filled as instructed by these:

```
#"Filled Down" = Table.FillDown(#"Renamed
Columns",{"Arrest"}),
#"Filled Up" = Table.FillUp(#"Filled Down",{"Domestic"}),
```

3. Fact Crime Data – several measures were created and used for reporting purposes.
   a. **Hour Alert 1**: hours were extracted using the time feature in the power query by adding a new column:

```
#"Inserted Hour" = Table.AddColumn(#"Changed Type1", "Hour",
each Time.Hour([Hour Alert 1]), Int64.Type)
```

   b. **Updated Hour**: similarly a column was created from the Date column when the system was updated:

```
#"Inserted Hour" = Table.AddColumn(#"Changed Type1", "Hour",
each Time.Hour([Hour Alert 1]), Int64.Type)
```

   c. **Hour**: Actual hours without minutes or seconds were created in a 24-hour format

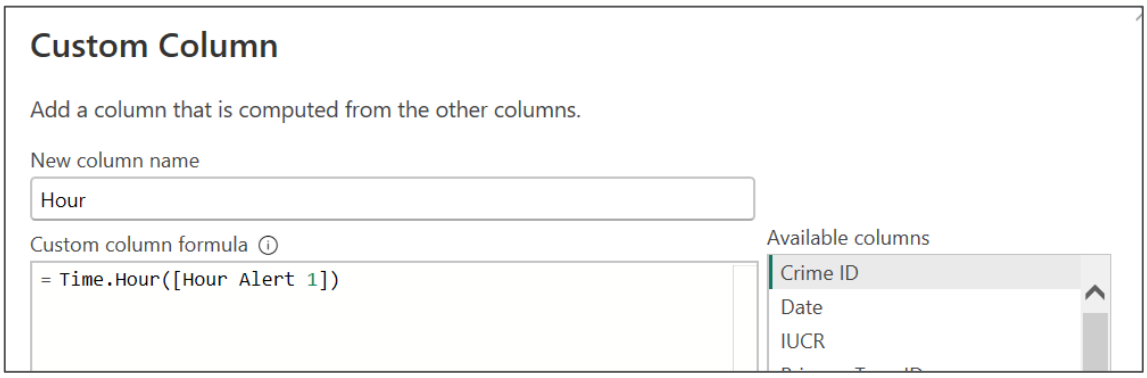


Figure 5: shows the custom hour column with the formula

   d. **Shift**: A conditional column to categorize the part of the day was created:

```
#"Shift - Conditional Column" = Table.AddColumn(#"Inserted
Hour", "Shift", each if [Hour] >= 15 then "Night" else if
[Hour] >= 7 then "Day" else if [Hour] >= 0 then "Morning" else
"Check")
```

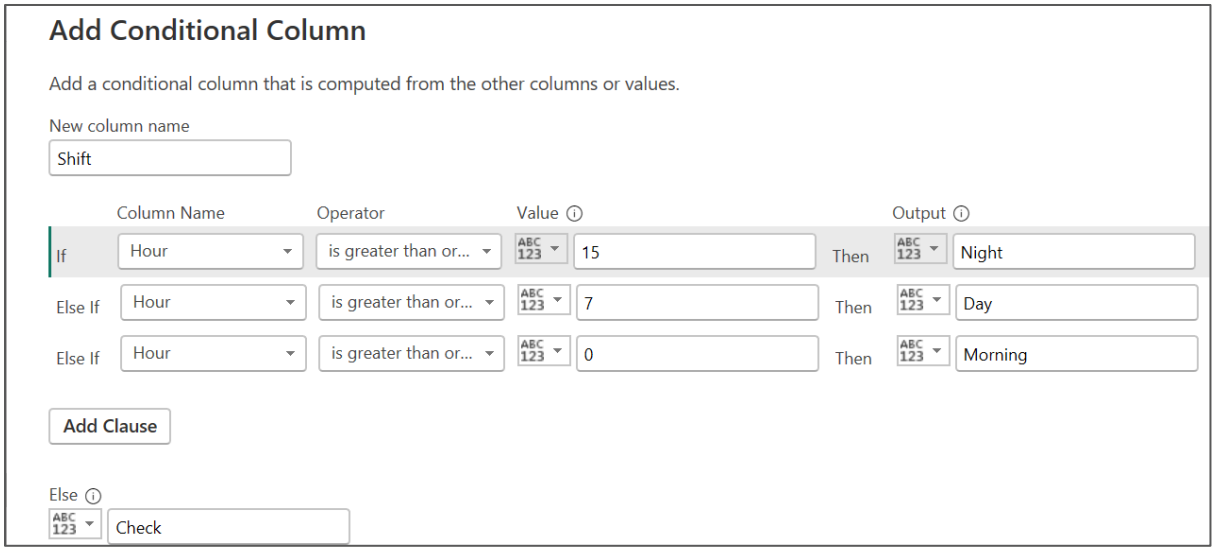


Figure 6: shows how we added a column to categories shifts.

9. MODEL OPTIMIZATION STEPS

For the model optimization in mind, various steps were implemented to enhance the overall performance and usability of the data model. These optimizations ensure a more efficient and organized data model, enabling better run time of the data model. Various transformations were applied to different dimensions, such as

- Standardizing data formats.
- Creating additional columns for improved analysis
- Removing irrelevant data.
- Unused columns are deleted where needed.
- Some unused columns are hidden which may be required for further development.
- Enable Load – been unchecked after a successful data load to **not reload** data each time when a change happens in the file.
- Only one background has been used throughout the entire visualization.
- No other color has been used for any buttons or cards.

By establishing hierarchies, we create logical relationships and levels of granularity within the data, enabling users to analyze and explore the information more effectively. For instance, we can implement a time hierarchy encompassing year, month, day, and hour, allowing Police to drill down into specific periods and gain insights at different levels of detail. Similarly, geographic hierarchies can be defined, such as city, district, and ward, facilitating spatial analysis and providing a hierarchical structure for examining data based on different geographic levels.

10. DESCRIPTION OF MEASURES AND CALCULATED COLUMNS

We have implemented various measures and calculated columns to address key analytical needs and provide valuable insights into our crime data. Firstly, we have created a measure that calculates the total number of crime events over time, allowing us to understand the overall crime patterns and trends over time. This measure can be sorted by the type of event, enabling us to analyze the distribution of different crime types. Theft and Battery are the top crimes in general whole years followed by criminal damage. There are 34 different types of crimes in 23 districts. For example, from 2018 to 2020 more than 529,000 crimes happened.

Additionally, we have incorporated a map visualization to identify the most affected districts and beats by crime events. By associating each event with a district and beat, we can analyze the geographical impact of crime. This map visualization complements a table that provides labels or names for each district, enhancing the interpretability of the data. The 001XX N STATE ST block is so dangerous because more than 2000 crimes happened in years.

Furthermore, we have analyzed to determine the average time gap between each occurrence of a crime event. This analysis helps us understand the temporal aspect of crime incidents and identify any patterns or trends in the timing of these events. We have also incorporated the ability to include different types of crime events in this analysis, allowing us to explore the relationship between crime types and time gaps.

To track the evolution of crime over time, we have included key performance indicators that measure changes in crime rates on a monthly and yearly basis. These KPIs provide a historical perspective and enable us to compare crime records across different periods.

Regarding data security, we have considered the importance of restricting access to relevant data domains. By creating security groups for each data domain, such as districts, we ensure that users have limited visibility of their respective district's information.

10.1. DAX MEASURES

To perform calculations and create derived data based on the existing data in your dataset we include DAX measures that allow us to take meaningful insights and answer specific business questions. These are the DAX measures that we created:

1. **Calendar table**: The calendar table was created using the code below:

```
Dim Calendar DAX =
VAR MinDate = MIN('Fact Crime Data'[Date])
VAR MaxDate = MAX('Fact Crime Data'[Date])
VAR CalendarDates = CALENDAR(MinDate, MaxDate)
RETURN
ADDCOLUMNS(CalendarDates,
        "Year", YEAR([Date]),
        "Month Number", MONTH([Date]),
        "Month Name", FORMAT([Date], "MMM", "en-US"),
        "Day of Week", WEEKDAY([Date]),
        "Day of Week Name", FORMAT([Date], "ddd", "en-US"))
```

2. **Total Crime:**
```
Total Crime = DISTINCTCOUNT('Fact Crime Data'[Crime ID])
```

3. **Total domestic crime event:**
```
Total domestic crime event = CALCULATE(
    [Total Crime],
    'Dim Crime Data'[Domestic]=True)
```

4. **Type Crime:**
```
Type Crime = DISTINCTCOUNT('Dim Primary Type Code'[Primary Type ID])
```

5. **Percentage of Arrested:**
```
% Arrested = DIVIDE([Total arrested],[Total Crime]/100)
```

6. **The number of Not Arrested:**
```
Not Arrested = CALCULATE(
    [Total Crime],
     'Dim Crime Data'[Arrest]=False)
```

7. **Total arrested:**
```
Total arrested = CALCULATE(
     [Total Crime],
     'Dim Crime Data'[Arrest]=True)
```

8. **Percentage of monthly evolution crime:**
```
% of monthly evolution crime = DIVIDE(
    [Total Crime],
    CALCULATE(
        [Total Crime],
        DATEADD('Dim Calendar DAX'[Date],-1,MONTH)),1)-1
```

9. **Percentage of yearly evolution crime:**
```
% of yearly evolution crime = DIVIDE(
  [Total Crime],
    CALCULATE(
        [Total Crime],
        DATEADD('Dim Calendar DAX'[Date],-1,YEAR)),1)-1
```

10. **Total hours of occurrence:**
```
Total hours of occurrence = SUMX(
   'Fact Crime Data',
    DATEDIFF('Fact Crime Data'[Hour Alert 1], 'Fact Crime
Data'[Updated Hour],HOUR))
```

11. **Average hours of the occurrence of a crime:**
```
average hours of occurrence of a crime = DIVIDE([Total hours of
occurrence],[Total Crime])
```

12. **Average crime per day:**
```
AVG crime per day = DIVIDE([Total Crime],
    DISTINCTCOUNT('Dim Calendar DAX'[Date]))
```
13. **Beats:**
```
Beats = DISTINCTCOUNT('Fact Crime Data'[Beat])
```

14. **Districts:**
```
Districts = DISTINCTCOUNT('Fact Crime Data'[Dist ID])
```

15. **Non-domestic crime:**
```
Non domestic crime = CALCULATE(
   [Total Crime],
    'Dim Crime Data'[Domestic]=FALSE)
```

16. **Reduction Crime Target:**
```
Reduction Crime Target = [Total Crime] * (1-[Value of What IF -
Reduction Crime TG ])
```

17. **Reduction Goal:**
```
Reduction Goal = ([Total Crime] - [Reduction Crime Target])
```

18. **Target:**
```
Target = CALCULATE([Total Crime] - [Total Crime] * 0.05)
```

19. **Target reduced:**
```
Target reduced = [Total Crime] -(([Total Crime] - [Reduction Crime
Target]))
```

20. **The year-over-year (YoY) change in total crime:**
```
Y-O-Y = DIVIDE([Total Crime],
CALCULATE(
     [Total Crime],
     SAMEPERIODLASTYEAR('Dim Calendar DAX'[Date])),1)-1
```

21. **Create a List of Values 1:**
```
[What IF - Reduction Crime TG ] = GENERATESERIES(0.02, 0.305, 0.005)
```
22. **Create a List of Values 2:**
```
[Value of What IF - Reduction Crime TG ] = SELECTEDVALUE('What IF -
Reduction Crime TG '[What IF - Reduction Crime TG ], "0.02")
```

23. **Create a List of Values 3:**
```
[What IF - Reduction Crime TG ] = GENERATESERIES(0.02, 0.305, 0.005)
```

10.2. DATA FORMATTING:

From all dimensional and fact tables:

- Duplicates were deleted from data tables.
- Columns were trimmed to reduce any spacing issues from all tables.
- Texts were capitalized for visual presentation.
- Dim Location: 0 and null values were replaced by “Unknown”. Details of the steps are as below:
- The date fields of all data files were changed with locale to avoid any inconsistency.

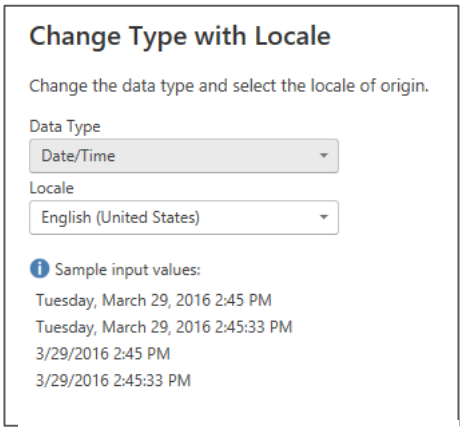


Figure 7: the date format we used for all dimension

11. ANALYSIS AND DISCUSSION

Through our comprehensive exploration of crime data, we have utilized various visualization methods to present the information effectively and gain valuable insights. Our visualizations cater to the specified project objectives, ensuring that they provide a clear understanding of the crime situation and support decision-making processes.

11.1. THREE MAIN KPIS

- **What if**

  Since we lacked target data for comparing performances across different dimensions and were unable to find a trustworthy external data source online, we developed a "what if" measure. This measure enables us to explore potential scenarios for various percentage-based key performance indicators (KPIs) within a threshold of 2% to 30%.

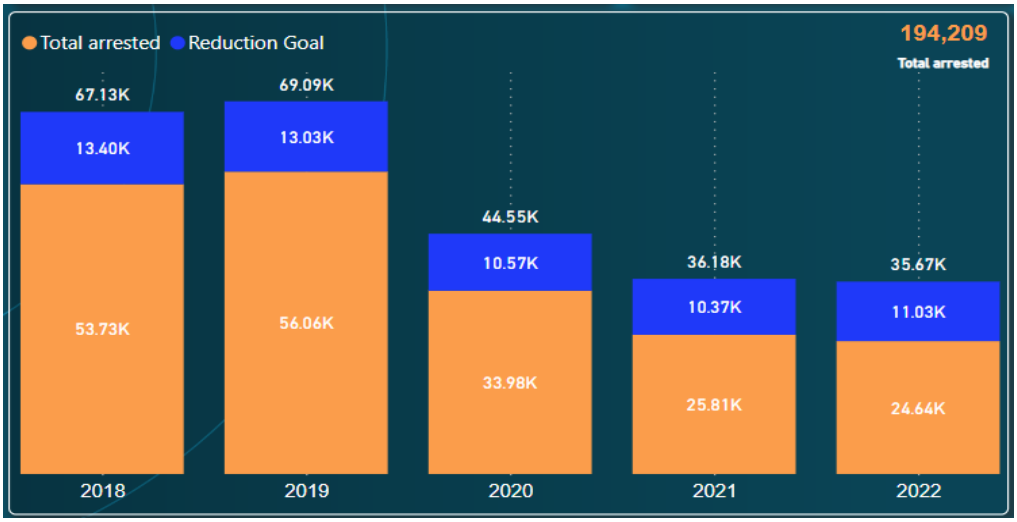


Figure 8: Shows the Impact of Percentage-Based KPIs on Arrest Rates

By utilizing the provided slide, users can observe how different percentages influence arrest rates.

- **Total Crime and Target**

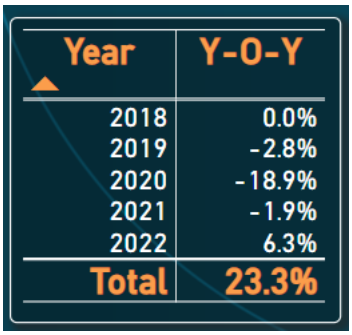
We have also developed another key performance indicator (KPI) specifically for the latest or selected year. This new KPI compares the total crime against the target crime. By utilizing this KPI, users can gain insights into the relationship between the actual total crime and the desired target crime for a given year.

Figure 9: Shows the total crime number and the target for a year.

- **A year-on-year comparison of total crime**

The third key performance indicator (KPI) focuses on the year-on-year comparison of total crime. This KPI provides users with valuable insights into the trends and changes in total crime over time.

By utilizing this KPI, users can easily observe and analyze the year-by-year increase or decrease in the total crime rate. This information allows stakeholders to identify patterns, spot significant shifts, and evaluate the effectiveness of crime prevention strategies implemented over different periods.



Year	Y-O-Y
2018	0.0%
2019	-2.8%
2020	-18.9%
2021	-1.9%
2022	6.3%
Total	23.3%

Figure 10: Year-on-Year Changes of crimes.

12. VISUALIZATION & MEETING PROJECT SCOPE

We have used a navigation bar throughout the project as it is more efficient and automated than using different buttons and bookmarks for navigation. It is also visually appealing on the published dashboard.

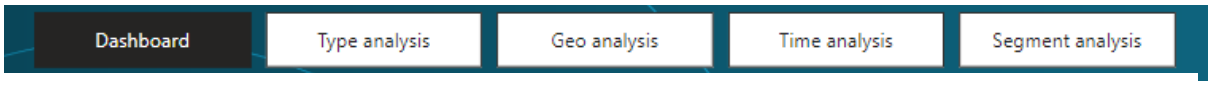


Figure 11: Navigation bar.

12.1. OVERVIEW

In the overview tab, the dashboard shows data on several areas of crime statistics throughout the year is presented. This view is divided into several areas, including the number of beats, the number of crimes that were committed, and the intended decrease in crime rates. A pie chart is also supplied to show how crimes are distributed among the three different shifts of morning, day, and night. To provide a complete picture of the crime situation, the goal Key Performance Indicator (KPI) and overall crime count are also presented. A bar graph contrasts the total number of arrests with the targeted reduction goal, while a percentage value shows the percentage of criminals who have been apprehended.

Figure 12: Comprehensive Overview of Crime Statistics and Law Enforcement Efforts

12.2. TIME INTELLIGENCE (TREND ANALYSIS 2018-2022)

One notable visualization is the yearly/monthly/daily/hourly crime events view, which presents the total number of crimes per day while allowing sorting by event type in a line chart. This visualization enables quick identification of crime trends and patterns, giving the Police of Chicago a comprehensive overview of the crime landscape.

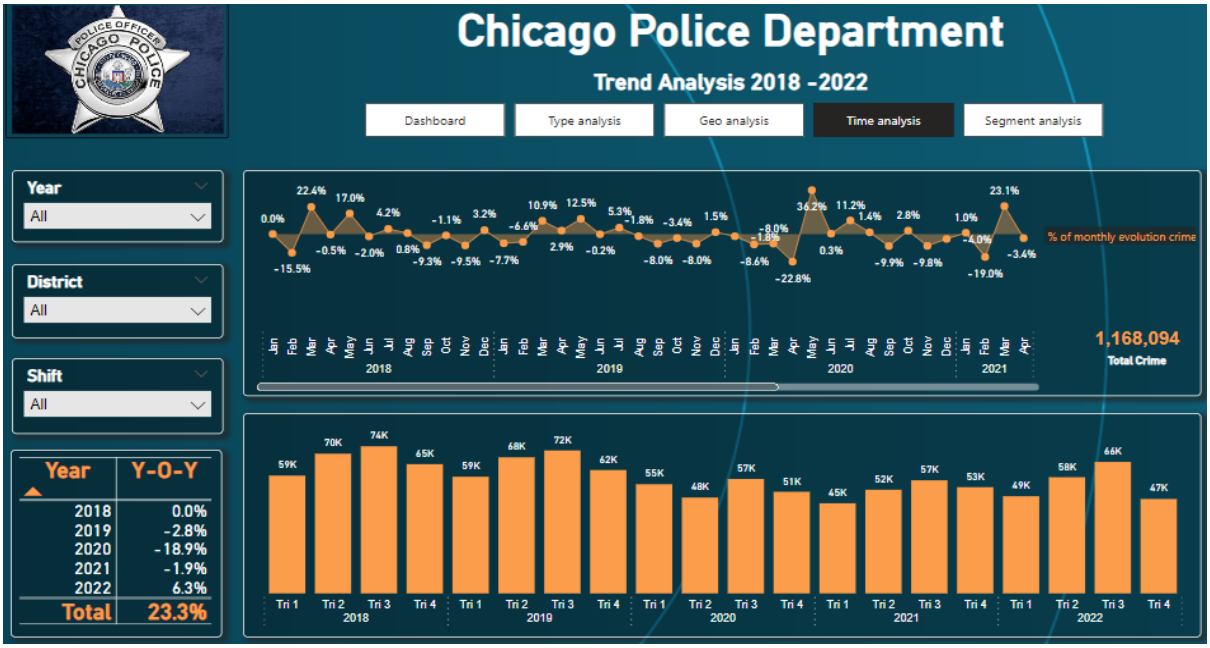


Figure 13: The number of crimes by year and month and year categorized by district.

12.3. GEOSPATIAL ANALYSIS

Another crucial visualization approach is the use of maps to showcase the impact of crime on specific districts and beats. By associating each crime event with its respective district and beat, we can visualize the geographical distribution of crime. This aids in identifying the areas most affected by crime and provides insights for resource allocation and targeted interventions. Furthermore, we have complemented the map visualization with a table displaying labels or names for each district, facilitating better comprehension and interpretation of the data.

Figure 14: The map of crimes in Chicago based on the district and can be filtered by year.

12.4. BY TYPE

To track the evolution of crime, we have developed key performance indicators (KPIs) that showcase the changes in crime rates over months and years. These bar charts, figure 6, provide a historical perspective and assist in comparing crime records across different periods, enabling Police to assess the effectiveness of crime prevention measures and formulate informed strategies. Information regarding the number of crimes, affected districts, beats, and type of crime. On this tab, we can have an overview of the evolution of crime from 2018 to the present day or filter the data by a specific year. In this way, the user can have access to information about a specific year, which makes the analysis more detailed. Bar plot considering data filtered by Type of crime. Each bar indicates the frequency of each type of crime whether the user selects all the year data or a specific year.

Figure 15: Type of the crimes

12.5. ADVANCED VISUALS WITH AI FEATURE

In PowerBI, we created two separate areas from this tab. The first one is "Key influencers," in which we filter data analysis considering a certain Day Name and visuals that support our inquiries so that we may gain insights into what can impact each Day Name. With the help of this tab, it is possible to comprehend the possibility of each Crime Type occurring on each Day Name, as well as the corresponding increase over the other Days. With the use of a bar plot, we can further analyze this data to determine the most prevalent Crime Type and the proportion of that Crime Type that occurred on the chosen Day Name.

Figure 16: Crime Segment Analysis based created by AI that when crimes increase or decrease.

A key influencers diagram has been implemented to visualize crime segmentations where a decrease and increase happen and possible scenarios.

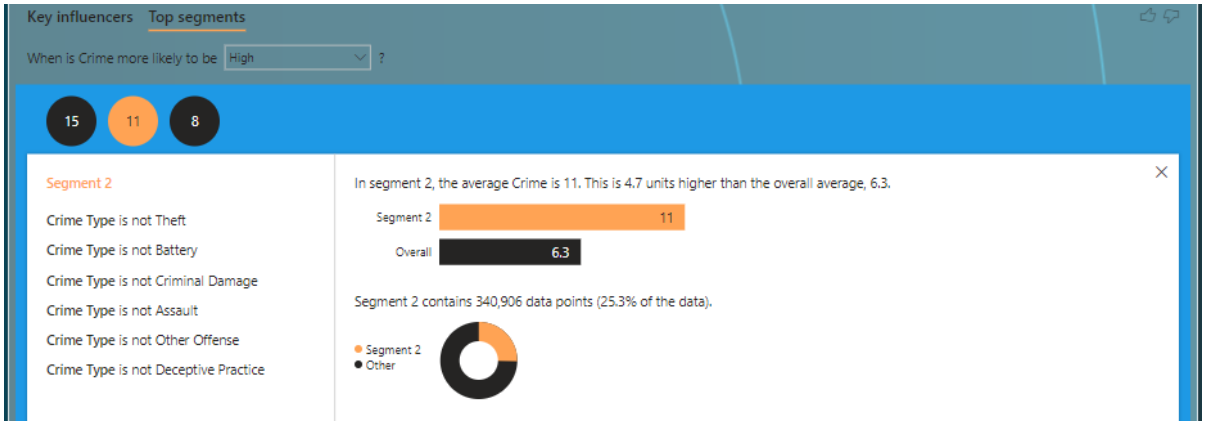


Figure 17: Shows one of the segments by crime type created by AI

13. ANALYSIS OF PROJECT OUTPUTS

- We successfully provide the total number of crime events per day and enable sorting by event type, allowing the Police to track crime frequency and identify prevalent crime types (accessible in PowerBI file).
- By analyzing districts and beats associated with each crime event and visualizing them on a map, we identify the most affected areas, empowering resource allocation and targeted interventions.
- Utilizing a map to depict crime hotspots and providing a table with district names or labels facilitates quick recognition of high-crime areas, enabling prompt action.
- Also, we tried to determine the average interval between criminal acts and accommodate different types of crimes in the analysis.
- Changes over time are recognized through comparisons of crime data across months and years, helping to identify prevailing trajectories and underlying factors, that bar chart can help us.
- Data visibility is restricted by sharing in PowerBI, ensuring privacy and confidentiality by limiting access to specific districts for respective users.

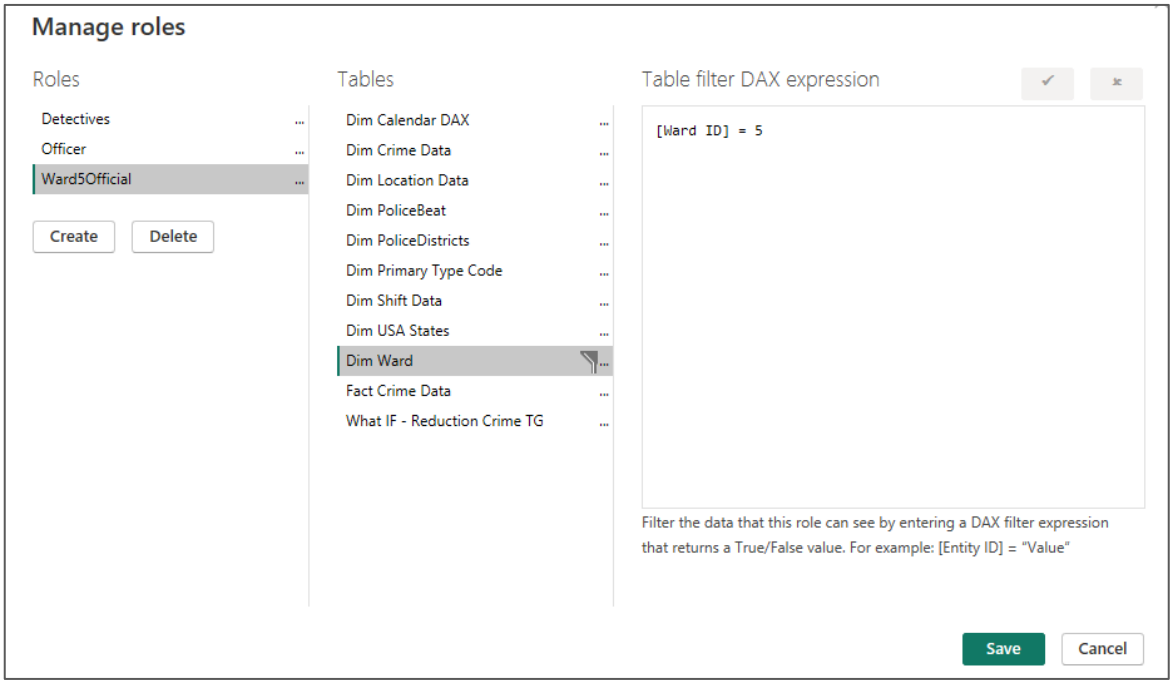


Figure 18: shows how permission can be managed for various personnel.

14. CONCLUSION

In conclusion, we can draw two aspects. Firstly, the technical aspect, where we can confirm that all the technical criteria have been met including the extra mile of setting up users' permissions roles and implementing a machine learning AI feature for in-depth analysis of data.

Secondly, regarding the client requirement aspect, this project provides the solution to all questions respectively. The model based on the PowerBI effectively meets the project's requirements by providing valuable insights into crime data. It includes specific variables such as "CrimeType" to categorize different types of crimes, "Arrested" to indicate whether an arrest was made, and "DomesticCrime" to identify domestic-related crimes. The model utilizes geospatial analysis, incorporating variables like "District" and "Beat" to identify the most affected areas. It tracks crime trends over time using variables such as "Year" and "Month," allowing for temporal analysis.

Furthermore, the model can be extended to analyze crime data in other cities and states, providing a comprehensive tool for law enforcement agencies to understand and address crime issues on a broader scale. By leveraging the power of predictive analytics, the model can help forecast potential crime hotspots, identify emerging patterns, and support proactive interventions.

In summary, the Crime Analytics data model not only facilitates a deeper understanding of crime data for the present but also offers the potential for future analysis and predictions, empowering law enforcement agencies with actionable insights to enhance public safety and security.

15. RESOURCES

Ferrari, A. and Russo, M. (2016) in Introducing Microsoft Power BI. Microsoft Press.

Ballard, C. (2006) “Dimensional modeling,” in Dimensional modeling: IN A business intelligence environment. IBM International Technical Support Organization, pp. 53–58.

Chicago Police Department (no date). Available at: https://home.chicagopolice.org/ (Accessed: April 15, 2023).

The population of Chicago, Illinois (city) (no date) The Demographic Statistical Atlas of the United States - Statistical Atlas. Available at: https://statisticalatlas.com/place/Illinois/Chicago/Population (Accessed: April 15, 2023).

Source of the header image: Creator: Kamil Krzaczynski, Credit: Getty Images, Copyright: 2021

2023 world population by country (live) (no date) 2023 World Population by Country (Live). Available at: http://www.worldpopulationreview.com/ (Accessed: 01 June 2023).

16. APPENDIX (LIST OF FIGURES AND TABLES)

Figure 1: This figure shows the relationships of tables in our PowerBI file

Figure 2: This figure shows the fact table properties.

Figure 3: The table shows the police shift categories.

Figure 4: the date format used for all dimensions.

Figure 5: shows the custom hour column with the formula

Figure 6: shows how we added a column to categories shifts.

Figure 7: the date format we used for all dimension

Figure 8: Shows the Impact of Percentage-Based KPIs on Arrest Rates

Figure 9: Shows the total crime number and the target for a year.

Figure 10: Year-on-Year Changes of crimes.

Figure 11: Navigation bar.

Figure 12: Comprehensive Overview of Crime Statistics and Law Enforcement Efforts

Figure 13: The number of crimes by year and month and year categorized by district.

Figure 14: The map of crimes in Chicago based on the district and can be filtered by year.

Figure 15: Type of the crimes

Figure 16: Crime Segment Analysis based created by AI that when crimes increase or decrease.

Figure 17: Shows one of the segments by crime type created by AI

Figure 18: shows how permission can be managed for various personnel.

Table 1: Describes the Data Engineering Steps by each table