ADTA 5250 - Large Data Visualization

Audience:

In my recent consultation with the city manager, a prominent figure entrusted with overseeing the welfare and safety of Dallas, I embarked on a journey into the realms of the Dallas Police Arrest data. Understanding the city manager's need for clarity in navigating the complexities of law enforcement data, the emphasis was on creating visualizations that are not only insightful but also accessible to someone with a moderate level of data literacy. This data was collected from May 2014 to November 2023 od Dallas County. Time data has also been collected, which can be useful to drill down day level analysis.

The city manager's mission is clear: to make informed decisions that enhance public safety, resource allocation, and overall governance. These visualizations aim to empower the city manager with actionable insights, fostering a data-driven approach to law enforcement in Dallas.

Information of the Data:

Location: Dallas County

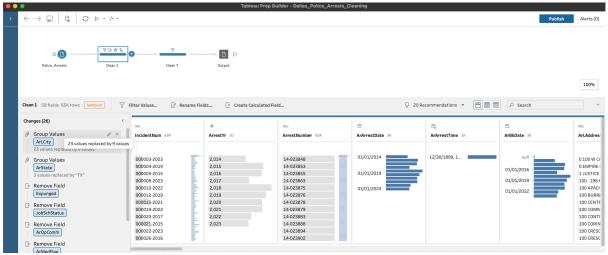
Premises: Highway, Street, Alley, Apartment Complex, Single Family, Gas or Service Station.

Weapons: Gun, Knife, Hands, Firearm, etc.

Quality of Data:

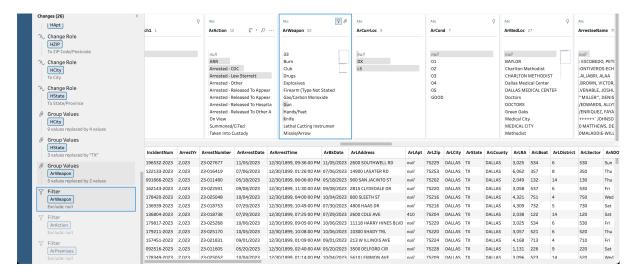
What all actions taken in this project,

- Removed rows with null values in critical fields such as ArrestDate, ArrestTime, and Location.
- o Applied a Group by filter for the "city" field to address consistency issues in city names.
- o Removed irrelevant columns or fields that do not contribute to the analysis.
- o Ensured consistent formats for date and time fields.
- o Checked for and eliminated duplicate records.
- o Documented changes made to the dataset, providing clarity on modifications.



I have noticed that there are spelling errors for the column city. For example, DALLAS has been misspelled as DALLLAS, DLS, DAL etc., ARLINGTON as ALRINGTON and many

other data values too. Hence I corrected them. There are 23 values has been replaced for this data cleaning. There are also few spelling errors in the State column as well, I have corrected them too.



I have identified null values in many columns in the dataset. May be there is a missing data in the dataset. Hence I removed null values using filter option in important columns like ArWeapon (Weapon Involved), ArAction (Action Taken) and ArPremises (Premises).

There are even more data cleaning process done on the dataset like converting data types or roles of city and state to geographical types, removed redundant or unused columns in the dataset since I have been not used in our analysis. This approach to data cleaning and maintenance enhances the overall quality and reliability of the Dallas Police Arrest dataset for analysis by the city manager.

Official Report for Dallas Police Crime Analysis

Overview:

Once upon a data-driven time in the city of Dallas, the city manager sought to unravel the story within the Dallas Police Arrest dataset. The dataset, a trove of temporal and locational details, became the canvas for our visualization journey.

The story began with a simple question: What tales does the Dallas Police Arrest dataset hold? The spark of curiosity ignited the first step – to explore and understand the nuances of the data.

On the morning of Thanksgiving, there was a brutal shooting at Dallas bar happened in the city of Dallas. It is breath taking to know that the shooting comes from any corner of the city. It is crucial to improve the law enforcement and tighten up the security concerns in the city especially at the times when the cases are more. So, in this analysis I will be illustrate Dallas city crime patterns, different cases registered at various locations and important actions taken to overcome them.

Our exploration into the dataset encompasses a series of strategic visualizations tailored to address specific questions and concerns:

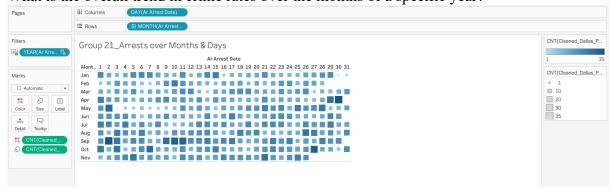
- 1. When are arrests most frequent throughout the day?
- 2. On which days of the week do I observe higher arrest rates?
- 3. How do arrest rates vary across different months of the year?
- 4. What is the overall trend in crime rates over the months of a specific year?
- 5. Which areas experience higher arrest counts, indicating potential hotspots?

In this process of analysis, there are numerous insights found with respect to crimes taken place in Dallas, I will systematically discuss them with proper measures and business findings.

Climax:

From our analysis, I found that answers for the predetermined questions about reckless crime arrests happening the city of Dallas. These answers have been formed from the analysis done on Tableau using the cleaned dataset.

1. What is the overall trend in crime rates over the months of a specific year?



In this calendar type view, I see the heat map correlated on the basis of number of arrests happened in particular day in the year. I can see overall picture of the arrests trend in this chart. Audience can easily understand which data has higher number of arrests. Hence I have put this chart.

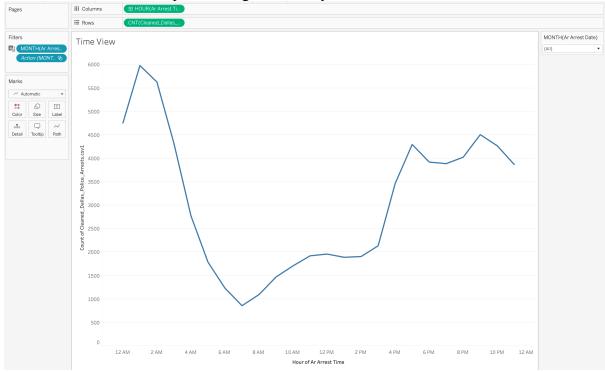
Business Insights:

- 1. From the chart, audience can see birds eye view of the arrests at every corner of the complete year selected.
- 2. In 2023, I see that September month have lots of hotspots days such as days 2, 9, 10, 11 have highest arrests count.
- 3. In April last week and May 1st week, I see there are higher number of arrests.

Unexpected Findings:

1. I see in the month of May 2023, there are very less count of arrests Hardly 1 and 2 Which is surprising to see such low number for multiple days. May be the data is not available.

2. When are arrests most frequent throughout the day?



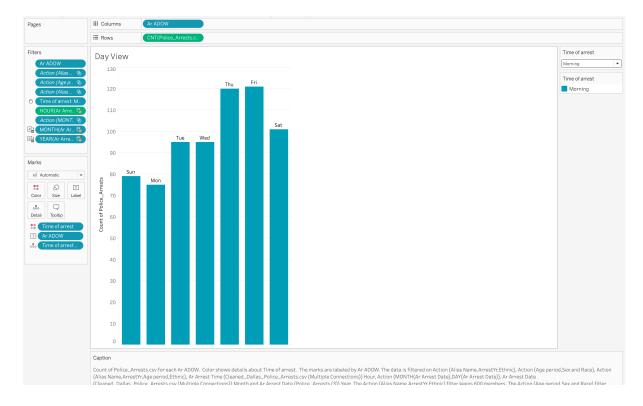
Stringent law enforcement cannot be in place all the day. Hence city manager and other audience can see this line chart to see which hour in the day have high number of arrests. Above chart shows number of arrests vs Hour of the day for the whole year 2023 but the same chart can be updated for any day based on the selected day on Month and Day View (heatmap).

Business Insights:

- 1. From the above, I see arrests are highest at 2 AM while lowest at 7 AM.
- 2. Briefly, I see that there are lowest number of arrests in the mornings and highest number of arrests in the nights.

Unexpected Findings:

- 1. It was unexpected that that the more number of arrests in the night after 12 midnight.
- 3. On which days of the week do we observe higher arrest rates?



From the above bar chart between Arrest count vs. Day of the week. Here, you can see which day of the week have highest number of arrests. I also used parameter to filter duration of the day such as morning, afternoon, evening and night In the above chart, the data is present for the complete year during mornings.

Business Insights:

- 1. I can understand that Thursday and Friday have highest number of arrests around 120, while Monday and Sunday have lower levels between 75 to 80.
- 2. I see that weekdays have highest number of arrests during afternoons.
- 3. During nights, you can observe that weekends have highest number of arrests.

Unexpected Findings:

- 1. It is surprising to see weekends are not as high as weekdays other than nights.
- 4. How do arrest rates vary across different months of the year?



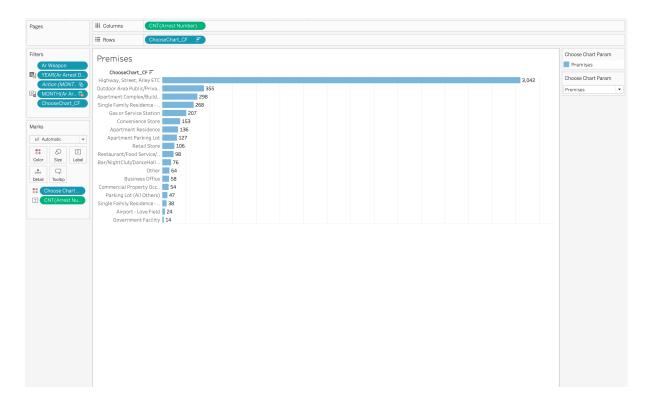
From the above line chart between arrest count and months in the year. This chart briefly explains what is actual trend in the year. In the above chart, I have filter which is selected as 2023, Hence the data shown is for the year 2023. You can see that September have highest arrests in the year 2023 with around 600 while February witnessed least with 360 arrests.

Business Insights:

- 1. You can see that September have highest arrests in the year 2023 with around 600.
- 2. February witnessed least with 360 arrests.

Unexpected Findings:

- 1. It is unexpected that February is the safest month in 2023.
- 5. Which areas experience higher arrest counts, indicating potential hotspots?



As the business decisioning charts I have plotted Premises wise, Action Taken and Weapon Involved in a single chart using parameter. Hence you can see Premises bar chart for the year 2023. From the chart, you can see higher difference between "Highway, Street, Alley" having highest count of arrests which is followed by "Outdoor Area Public/Private". These categories changes based on the selected filters.

Business Insights:

- 1. From the chart, you can see higher difference between "Highway, Street, Alley" having highest count of arrests might be due to speeding, accidents etc across roads. which is followed by "Outdoor Area Public/Private".
- 2. I also see that "bar, pub and other" premises have higher number of arrests.

Unexpected findings:

1. There are many arrests in the "apartment complex" and "single family residence".

Ending:

- 1. From the analysis, I recommend that law enforcement must be made strongly at nights during weekends especially near bar, pubs and public places. So that peaceful environment will be kept.
- 2. There are many arrests took place during holidays over the years, so I need special restrictive and better security needed during holidays
- 3. Highways have highest number of arrests, number of speeding cameras must be increased so that I can mitigate falseful activities across city.