**ADTA-5250: Assignment - Creating Dashboard**

**Sri Charan Bodduna - 11661666**

**About Data**

Dallas Police Arrests dataset containing arrests made by police in and around the city of Dallas. There are 65 columns and around 102653 rows. Each row denotes an arrest that took place in and around the city of Dallas. There are different sets of columns present in the dataset. There are a set of columns representing details of arrests like police personnel name, time, place, and weapon found. There is demographic information about suspects like age, weight, race and employment status, few columns about address. The data is from 2014 to 2023.

**Data Cleaning**

A screenshot of a computer

Description automatically generated

Fig. I: Data Cleaning changes

As part of the data cleaning step, I have done a couple of cleaning in the data which is required. We can see the changes in the above screenshot. Few of them include, removing null values, correcting city names and state names, changing the data types of city and county to geographic types. There are few columns which are redundant in the analysis, hence I filtered them out. All these cleaning steps are performed in Tableau DataPrep.

**Dashboard**

A screenshot of a computer

Description automatically generated

Fig. II: Dashboard

In figure II, I have created a dashboard showing various charts and Key Performance Indicators (KPIs). There are also filters and calculated fields in the respective charts. The above dashboard is interactive with the help of county filters and map as a filter (when we select any county on the map based on the selection other charts like pie, line and bar chart get updated respectively).  There is another filter for Premises. I have chosen Premises as a filter so that we can clearly understand in depth view of the arrests taking place majorly. Later we see a pie chart for interpreting the composition of different kinds of weapons involved in the arrests. We see the top 6 kinds of weapons in any selected county. Finally, we see the arrests pattern across years using line charts, we see clear trends of arrests using line charts. I have also annotated highest cases and lowest cases occurred for each county in the line chart.

I have added a descriptive title, my last name in the charts title, a logo and followed gestalt principles and ensured pre-attentive attributes with color and size.

**Business Insights**

1. From the dashboard, I understand that Dallas County has a huge number of arrests. So, It is important to tighten the security and public safety measures in the city of Dallas with critical places like Highways and Outdoor areas.

A screenshot of a computer

Description automatically generated

Figure III: Downward trend in arrests in 2023 in Collin County

A screenshot of a computer

Description automatically generated

Figure IV: Downward trend in arrests in 2023 in Dallas County

1. I focus on the recent trend of arrests happening in all counties. With my filter on the county, I have understood that the arrests count has been reduced in the year 2023 from the previous year. As in the above figures of the line chart with filter selection for Dallas and Collin counties but it is the same scenario with Denton as well.

A screenshot of a computer

Description automatically generated

Figure V: Highly critical Premises in Denton County

1. I wanted to understand the critical location in Denton County so that I can be careful. For knowing that, I have used filters and selected Denton and highly critical locations Premises. With that, I got to know Highways, Apartment Complex and Parking lot are the places where arrests took place largely.

**Sheet 1:**

A pie chart with different colored circles

Description automatically generated

Figure VI: Pie chart for weapon proportion

In this pie chart, I have used the Weapon column in the dataset and categorized on the same column and performed the count of arrests so that we see which weapon is hugely involved in the arrests. Above screenshot shows the data of all counts with top 6 weapons.

I have chosen a pie chart to understand the clear proportion of the weapons and number of arrests.

**Sheet 2:**

A screenshot of a computer

Description automatically generated

Figure VII: Bar chart for Premises

In this bar chart, I have used the Premises column in the dataset and categorized over different Counties and performed the number of arrests that took place so that we see which Premises or areas are having a greater number of arrests. Above screenshot shows the data of all counts with top 6 weapons.

I have chosen a bar chart to interpret the clear difference in the Premises and the number of arrests. We see here the highest arrests of 32463 took place at the highways, Streets and Alleys.

**Sheet 3:**

A screenshot of a map

Description automatically generated

Figure VIII: Map view of the County

In this chart, I have plotted a map chart for the different counties. I have color coded each county on the map. This chart is helpful when we want to see birds eye view of the arrests.

Most important thing here in this chart is that I have used this chart as one of the filters for the data in the dashboard. So, we can easily select the county and see different measures about it.

**Sheet 4:**

A graph on a screen

Description automatically generated

Figure IX: Line chart for the arrests over time

In this line chart, I wanted to see the arrest patterns over the years from 2014 to 2023. So, I have taken the arrest date column from the dataset and taken the arrest count on the y axis. I have also labeled line charts with highest and lowest arrests found. So, for example, the highest arrests in Denton is 96 in the year 2017 and least is 32 in 2023. This makes the audience easily interpret the significant years with low and higher rates of arrests.

I have chosen a line chart to understand the trend and patterns of the arrest count over the span of time.