**Assignment: Filters and Parameters**

Sri Charan Bodduna – 11661666

**About Dataset:**

Dallas Police Arrests datasets contains arrests made by police in and around the city of Dallas. There are 65 columns and around 103000 rows. Each row denotes arrest took place. There are different set 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 suspect like age, weight, race and employment status, few columns about address. The data is from 2014 to last month (October 2023).

**Chart 1:**

A screenshot of a computer

Description automatically generated

Figure 1

In the first place, my question is about identify number of arrests happened majorly due to which weapon? I have plotted a good-looking color-coded bubble chart with number of arrests represented by size of bubble and color denotes type of weapon involved in the arrest. I have used filters to get the weapon involved in highest arrests. I have also used parameters for weapon for selecting and arrest count to shorten my measurements and limit what I am looking at.

**Business insight 1:** I see that highest number of weapons found in the arrests is of “Handgun”. I felt astonished! The number of Handgun arrests across different counties and cities is 3520. Which is a questionable number which we need to be addressed. This insight can be very useful to take informed decision about gun usage at places.

**Business insight 2:** From the data over years, we see in the above plot that top four weapons involved are Handgun, gun, Hand/feet, Pocket Knife. With these weapon types, there are highly dangerous weapons stood at the top of the chart. These top weapons constitute 8933 arrests among total arrests.

A screenshot of a computer screen

Description automatically generated

Fig. 2

**Insight 3:** I am interested to know which weapon are found to be least in the arrests. Hence, I used filters to drill down into the data and select only weapons with least arrests count. We can see in above figure that Burn, Arrow, Explosives and Strangulation. Which means the there are no major weapons involved with least counts.

**Chart 2:**

A screenshot of a computer

Description automatically generated

Fig. 3

In this analysis, I wanted to know which places/premises of the cities have highest number of arrests. Since my data is huge and biased with Dallas, I have taken the sample of Denton and Collin counties by using filters and parameters. With that I have plotted stacked horizontal bar chart to get insights. Also, my choosing “Ar Premises” filter I have taken places with higher arrests (ignore least numbers).

Below are the insights from the chart:

Insight 1: In most of the places/premises, we see that Collin County have leading numbers than Denton. Which tells me that Denton is safer than Collin in this context. Which is good and I feel safe in city of Denton.

Insight 2: While looking above chart, we see that premises like highway, Apartment Complex, Apartment Parking lot have greater numbers when compared to others. Which means we need to be cautious while traveling over highways and staying at home during nights. This rises an important data point to the public safety officials to keep an eye in these premises vigilantly.

Insight 3:

From the above chart, we also look that few premises are having higher arrests in Denton. For example, Retail Store, Convenient Store, Clubs and Bars. Interestingly, we see that only different kinds of stores are higher in number in Denton. This leads an important discussion to the city municipal officials and public safety department to keep surveillance at the public and minimize arrests with awareness and personality development programs.

I have also created Parameter in the chart.