Information Visualization Final Project - Austin Crimes

Haider Ahmed (ha1065)

13th December 2019

Abstract

For this final project, I will present a Crime data visualization system for the Austin. The system will allow potential new residents and security organizations the ability to explore the data and to take advantage of the analysis presented. After conducting data analysis, I build a system to visualize: 1)The number of crimes and crime rate of particular council areas within the city; 2)Solved and unsolved crimes based upon number and dates 3)how long it took for a particular type of crime to be solved in particular areas of Austin. Through visualization, the study provides a deep understanding of the crime prevalent in Austin in 2015.

1 Introduction

In Austin, TX you have a 1 in 242 chance of becoming a victim of violent crime. Violent crimes include murder, rape, robbery and assault. With regards to property crime, you have a 1 in 32 chance of becoming a victim. Property crimes include burglary, theft and vehicle theft. Overall, you have a 1 in 28 chance of becoming a victim of crime in Austin, TX. All of these statistics are meaningless, if the data behind them isn't analyzed to highlight patterns in locations and time to more effectively carry out preventive actions. Therefore there is plenty that can be achieved through a visualization of the data, problems such as 1) Which areas of Austin have the most crime ? 2) What type of crimes are prevalent in various localities of Austin 3) What is the Austin PD's performance in solving these crimes ? can be understood on a deeper level; and it is only when a problem is understood that measures can then be taken to rectify it.

2 Data

The data that I use is the official Austin, Texas government's Annual Crime Dataset for 2015, which is available at : https://data.austintexas.gov/Public-Safety/Annual-Crime-Dataset-2015/spbg-9v94. This dataset encompasses all of the Part 1 crimes in Austin, Texas from January 1 2015 - December 31 2015. There are a total of 38.6K data points with 13 features each.Each crime is identified by a unique 'Go Primary Key'. Apart from that there are columns such as 'Council District' and 'Go Location Zip' which specify the zip code and council number of where a particular crime took place in the city. There are columns which identify the crime itself, the crimes included are 'Auto Theft', 'Robbery', Burglary', 'Theft', 'Assault', 'Rape' and 'Murder'. There are fields which identify the status of the crime , whether it has been solved or not and the date they were solved if they were. The remaining features are not integral to the aim of my visualization, therefore I will not mention them.

2.1 Preprocessing

My first step before even attempting visualizing the data was to process the raw data so that I have a dataset which will make it easier to achieve my visualization goals. As the major focus of my visualization would be to discriminate between solved and unsolved crimes and provide a rudimentary analysis of them , I had to use Python to first remove about 2000 data points, corresponding to crimes which have

no 'Clearance Status'. Furthermore, than I created two additional columns with in the Dataset, mainly 'Report Day' and 'Clearance Day' which would represent a day of the year 2015. For example, if a crime was reported on 1-Jan-2015, that would correspond to an entry of 0 in my 'Report Day' column and similarly for my 'Clearance Day' column.

2.2 Flask and MongoDB

The processed dataset was then imported into a MongoDB database. The reason I decided to do this was because I felt it would be much easier to manipulate and filter the data through queries I make to the database rather than writing javascript code to do the same. I then used flask to create a web application for my visualization.

3 Visualization Design and techniques

1. An Interactive Map of Austin was created with the help of d3.map and using a GeoJSON dataset of Austin's 10 Councils,using Mercator Projection. The map presents the geographic information of the city of Austin and provides a high level view of the crime statistics in each 'Council'. The color of each council represents the Crime Rate of that particular council. The darker the color, the more crimes took place in this council over the course of 2015. Hovering your mouse over a certain council shows the Council District Number , Crime Rate and Crime count of that council district. The councils with the highest crime rate are Council District 9 (17%) and Council District 3 and 4 (15% each). Overall speaking, the crime rate of western Austin is on average higher than that of the east. This map serves as the root of the rest of the visualizations, meaning that selecting a certain council area on the map, will filter the rest of the visualizations with the relevant data.

2. An interactive brushable line chart (Figure 2) is implemented which displays crime occurrences of each crime for each month over the 12 months of 2015.These are color coded by crime types. This timeline chart brings out time/seasonal relationships

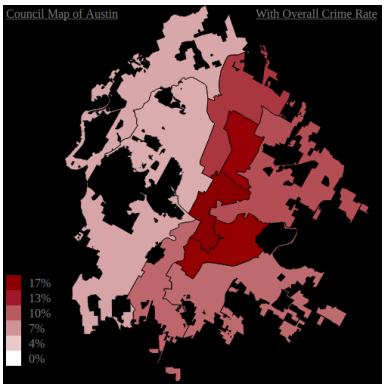

Figure 1

between different crimes. An additional feature that this chart possess is the ability to zoom in. That is selecting a range expands the chart to focus more on the selected region.

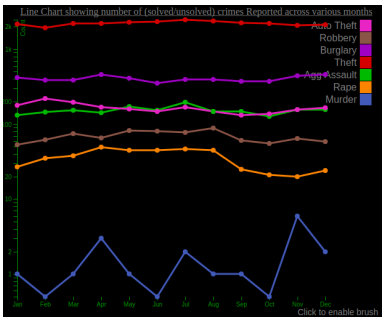

Figure 2

3. A grouped Bar graph(Figure 3) is also implemented which gives information about the number of solved and/or unsolved crimes reported in a chosen council. Red represents unsolved crimes and green represents solved crimes. Each bar gets highlighted on mouse hover, and displays the crime type , whether its represents solved or unsolved crimes and the crime count.

4. A Bubble chart (Figure 4) was also made using the capabilities of D3.js. The purpose of the bubble chart was to show how many crimes were committed at each location which is determined by the ZIP code. The Bubbles are color coded per council. This

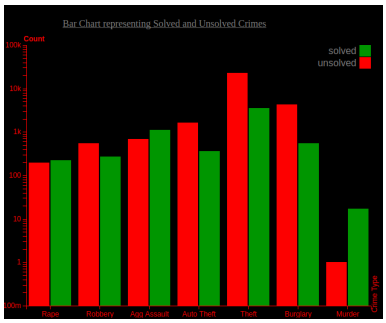

Figure 3

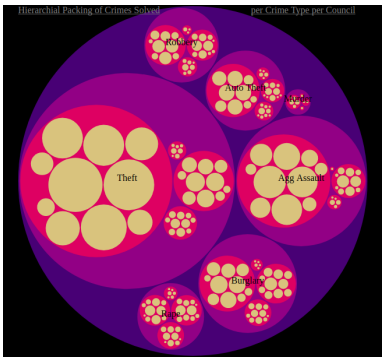

Figure 5

is useful for new residents to Austin, who could determine which area they would like to move in depending on the number of crimes. A view like this could also prove useful for police authorities in determining which areas to increase their presence and oversight.

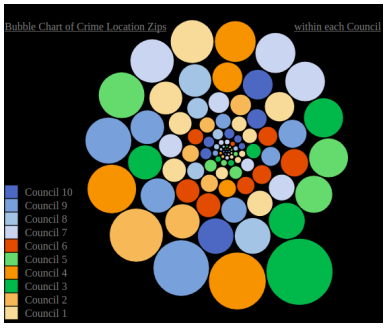

Figure 4

5. The final chart that was plotted is the Hierarchical Circle Packing (Figure 5). This proved to be one of the most insightful charts that one can obtain through the capabilities of D3. The reason being the depth of information that can be attained using this chart. Using this chart the information about the various types of crimes committed can be attained at great lengths. It can be classified based on the councils and also can give information on the number of days it took to solve the crime.

4 Conclusion and Future Work

Based on the five visualization designs, we have a deeper understanding of the crime prevalent in Austin and the underlying patterns connected to certain crime types and in certain areas . It is true that visualization provides a more effective and efficient way to convey information compared with showing pure numbers. In the future, there could be many improvements if more time are allowed. Firstly, a much needed improvement would be to add a more up-to-date dataset. Our current data is only limited to one year, 2015, which is 5 years in the past. It is essential that contemporary data is fed into the visualization system so that we not only have a better understanding of the crime happening today rather we can also gain valuable insights by having visualization designs that compare various types of crimes through out various years.

5 References

- Austin Police Department. "Annual Crime Dataset 2015: Open Data: City of Austin Texas." Data.austintexas.gov, 28 Nov. 2016, https://data.austintexas.gov/Public-Safety/Annual-Crime-Dataset-2015/spbg-9v94.

- Areavibes. "Austin, TX Crime." Austin, TX Crime Rates Crime Map, https://www.areavibes.com/austin-tx/crime/.

- D3. "Zoomable Circle Packing." Observable, Observable, 29 Oct. 2019, https://observablehq.com/@d3/zoomable-circle-packing.
- "Interactive Data Visualization with D3.Js, DC.js, Python, and MongoDB." Interactive Data Visualization with D3.Js, DC.js, Python, and MongoDB // Adil Moujahid // Data Analytics and More, http://adilmoujahid.com/posts/2015/01/interactive-data-visualization-d3-dc-python-mongodb/.