A Chicago Crime Analysis Using A Visual Approach

David M. Berenstein
Eindhoven University of Technology *

Mike de Brouwer
Eindhoven University of Technology †

ABSTRACT

This paper dives into the vast amount of data publicly available on the Chicago data portal, in particular we look at crime statistics. This data set contains every crime registered in Chicago since 2001, including information such as location, date and type of crime. Using this data multiple insights can be gained, we identified four important questions that will help us find out where and why crime is happening in Chicago. These questions are answered by aggregating the data in various ways and using interactive visualization techniques to allow the user to gain insights to help answer these questions.

Index Terms: Bubblemaps—Interactive—Visualization techniques—Chicago Crime Analysis—Scatterplotmap—Design Principles—Design Perception

1 INTRODUCTION

With the rise of the complexity and amount of data, it becomes more difficult to gain a proper understanding of the significance of data. By applying proper visualization techniques it is possible to get a better understanding of this significance. Visual encoding for information visualization, interaction, multiple views, and the exploratory analysis process are part of the process of visualizing a data set.

The city of Chicago has a data portal that allows members of the public access to a wide variety of data. One of these data sets contains all crimes that happened in Chicago since 2001. This is a massive set of data and the challenge is to apply various techniques to present this data in a comprehensive and understandable way such that insights can be gained that might help Chicago combat these crimes.

This paper is structured as follows, section 2 elaborates on the problem description and task analysis. Section 3 and 4 go over the visualization techniques used and their use cases. Section 5 discusses the approach used and section 6 concludes our findings. Finally section 7 shows a breakdown of the work done.

2 PROBLEM DESCRIPTION AND TASK ANALYSIS

The city of Chicago has a large structured data storage[3] and wants to stimulate gaining insight into this data. This data storage contains a data set that covers crime-related data from the years 2001 till 2018. Each row in this data set resembles a specific crime.

For each of these crimes the following column information is considered relevant: 'Date', 'Primary Type', 'Description', 'Location Description', 'Arrest', 'Domestic', 'Beat', 'District', 'Ward', 'Community Area', 'Latitude', 'Longitude'. The Primary Type, 'Description' and 'Location Description' contain information about the type of crime like 'THEFT' with a description like 'under 500 dollars' and the location were the crime occurred like 'STREET'. Arrest and Domestic are boolean values that indicate whether the crime resulted in an arrest or was and occurred in a domestic relationship. Lastly, Beat, District, Ward, Community Area are integers that refer to a specific region in the city of Chicago.

We also found some options for additional data sets. Because no other consistent yearly sources could be found we decided to stay with the other data sets offered on the data portal of the city of Chicago. Most data sets did not provide much information on their own but could be a good addition to the explain behaviours in the crime data.

The police in Chicago has had difficulties with gaining insight into the crimes that are committed. In order to help the police, this paper shall evaluate and apply four separate visualization techniques that take perception and design principles into account. Here perception and design principles are considered such as scale distortion and chart junk by keeping the visualizations clean and in their appropriate contexts. We are not planning to use 3D, given that our data set is quite big our goal is to use aggregation and filtering steps to make it more understandable. When possible we will try to limit ourselves in the amount of different colors used to keep the overview.

By applying these visualization techniques this paper provides several tools that help persons interested in these crimes to gain insight-based on four research questions:

- What causes the different crimes in different areas of the city?
- Where are specific crime types located?
- When do specific crime types occur?
- How do crimes spread throughout the city?

Section 3 shall first introduce each of the visualization techniques and shall then provide insight into the implementation of each of the techniques.

3 VISUALIZATION DESIGN

3.1 Evaluation of Methods

This section focuses on the evaluation of visualization methods. The visualization methods are from two different Python libraries that are similar to the visualization package D3 for Javascript[2]. Plotly[6] and Bokeh[7] provide standard examples of common visualization techniques. These visualization techniques are mostly implemented in an IPython Notebook. Such a notebook can then be opened and edited inline using Jupyter Notebook.

Jupyter Notebook has a standard framework for implementing interactive widgets like buttons and sliders. These widgets are called the IPyWidgets. The specific choice for these widgets shall be discussed in sections 3.2, 3.3, 3.4 and 3.5 because these cover the implementations the chosen visualization techniques.

Plotly distinguishes seven chart types: basic, statistical, scientific, financial, maps, Seaborn, 3D. Because we do not deal with statistical, scientific, 3D or financial data these visualization types are not considered to be an option for implementation.

*e-mail: d.m..berenstein@student.tue.nl
†e-mail: m.a.j.d.brouwer@student.tue.nl

The first chart type we consider is the basic charts. The basic charts are simple bar charts, histograms, line plots and scatter plots and do not offer many interactive visual insights by themselves. Given the fact that we are dealing with location data it might be interesting to combine a scatter plot with a map. In addition, we could consider using an interactive bubble chart/line to get a general overview of the crime evolution over time. Especially the bubble chart might be used for implementation with an additional data set.

The second chart type considered is maps. Map visualizations were created for visualizing geographical data. Whereas the scatter plot has the potential to point out specific crimes, you can also implement a map-based on aggregating-based on geographic regions or neighbourhoods. Because our data has information about in which Beat, District, Ward or Community Area the crime was committed it might also be interesting to aggregate the data-based on these regions. The shapes for these regions can again be derived from the data portal of the city of Chicago.

The third chart type considered is seaborn maps. These maps are partially scientific, but also offer density, heat map and 2D histogram overviews. Because the police want to gain insight into the moment during which crimes are committed we also consider making a time-based heat map per area per crime. Using this the police can get an overview of hot periods during which certain crimes are more likely to occur.

For the consideration of visualization techniques, Bokeh did not offer any new techniques. However, Bokeh did offer some additional examples on how to implement interactivity in plots.

As a result of these considerations we choose to implement an interactive geographic scatter plot, a bubble chart animation, an interactive choropleth map plot and time-based crime heat map.

3.2 Interactive Scatter Plot

A scatter plot is a figure that shows a data entry based on coordinates on a map. Because the data set contains coordinates of each crime that was committed, we decided to make a geographic scatter plot. This scatter plot had been implemented using the 'scattermapbox' function within plotly.

The scattermapbox function scatters each data point on a Google Maps-like geographic mapping environment based on their 'latitude' and 'longitude'. Each point is assigned a colour based on categorization. We decided to apply a categorization based on the primary crime type. In addition, you get a hover function with the location of the specific plot. Figure 1 shows this implementation. Even though this standardized implementation functions it is a basic implementation and suffers from two problems.

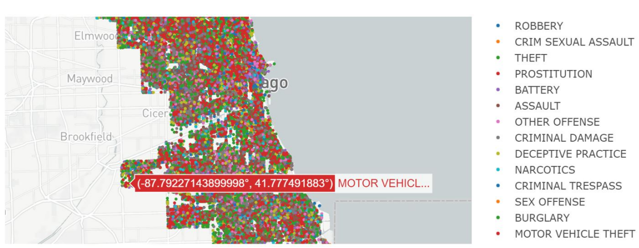

Figure 1: Standard Scatter Plot

In order to improve the functionality and readability of the scatter plot, we altered the hover function and the size of the plot. The coordinates do not add any information to the hover and the hover is very wide. We chose to let hover function show 'Date', 'Primary Type', 'Description', 'Location', 'Arrest' and 'Domestic' information for the crime. Furthermore a dropdown menu is added to alternate between different map types. Figure 2 shows this implementation.

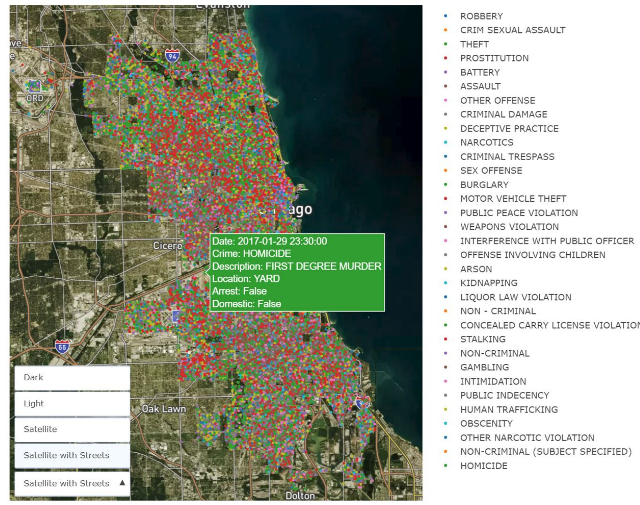

Figure 2: Improved Scatter Plot

The last addition to this scatter plot is adding interactivity and filtering of the data. This filtering is not a standard functionality and was achieved by first adding multiple IPythonWidgets. Then these widgets are linked to a scatter updating function, which relocates the scatters. Because it is not possible to delete or create scatters after the plot is made we chose to implement this workaround via the updating function. These widgets provided selection criteria for an hour, day, weekday, month, year, 'Primary Type', 'Location', 'Arrest' and 'Domestic'. In addition, we added some accordion boxes to hide the functions in. The time-related filters are used via two-way sliders and the rest are used via selection boxes. Lastly, a button is added to update the figure based on the filters. Figure 3 shows the interactive scatter plot filtered on homicides that occurred in houses in November 2016. Note that only a part of the actual map is shown. Actual use cases will be discussed in section 4.

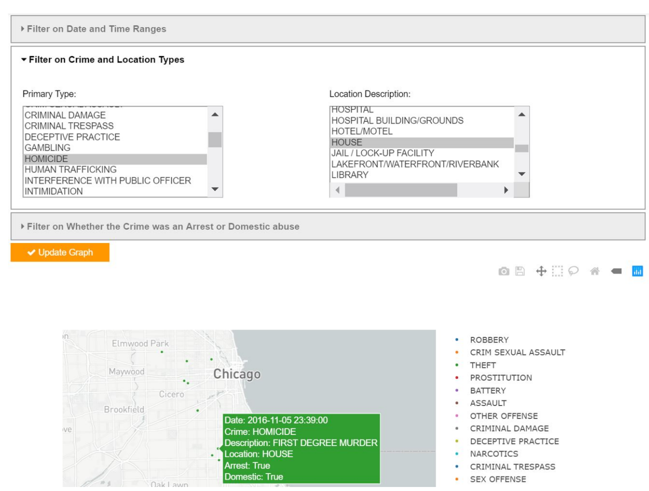

Figure 3: Improved Scatter Plot

3.3 Bubble chart Animation

A bubble chart is very use full when you wish to plot many different variables in a manner that still provides a good overview. This is due to the fact that besides a simply x and y axis with data you can also convey more information through the size and color of the plotted data.

The focus of this visualization is the question 'What causes the different crimes in different areas of the city'. For this purpose the data set containing crimes in Chicago is extended using a data set from the same data portal [1]. This data set contains various socioeconomic indicators for each community area in Chicago, these are: percent aged 16+ unemployed, percent households below poverty, percent aged 25+ without a high school diploma, percent of housing crowded, percent aged under 18 or over 64, per capita income and finally a hardship index which is a number between 1 and 100 that indicates hardship based on the previous indicators.

The data is aggregated over the period 2008-2012, for this reason the first thing we do is to only select the crime data in that period as well. Because we are interested in what causes the different sorts of crimes we split up our data set into the different primary types of crimes we have such as robbery and homicide. Then finally we aggregate the data per community area, thus we count how many crimes of a certain type have been committed in each community area over the period 2008-2012.

So now we have quite a few variables that we would like to visualize, yet even a bubble chart can not visualize them all at the same time. Since it is easy to animate a bubble chart by going over the years we first thought of doing that. But given that the data about socioeconomic indicators is aggregated over 5 years we decided against it. While we could go over the crime data year by year this would not tell us anything since the indicators would not be accurate for that. So we first started with a simple singular bubble chart to see how the data would look like and tried out various options for the axes, an example is shown in figure 4. Each bubble is a community area with the size of the bubble being the hardship index, the hardship index incorporates all indicators so we felt it should be prominent in the graph. The y-axis counts the number of crimes, and the x-axis shows one of the indicators. We tried out various options for the color coding and we settled on the per capita income because that usually ties in nicely with the other indicators such as unemployment rates.

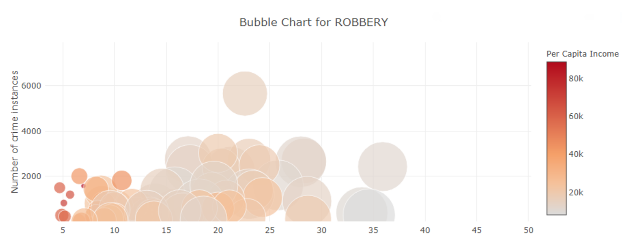

Figure 4: Bubble chart first version

Now there are still some issues with this chart, it currently only shows two of the indicators, only shows one crime type and the results get skewed by not accounting for population. Luckily we were able to find census data per community for 2010 and we normalized the crime rate for population. Now the idea is to allow for comparisons between crimes and indicators by plotting two charts and giving the user the ability to change what crimes and indicators are used. Figure 5 shows the controls that changes the charts in figure 6. You are able to change what crimes are shown in the top and bottom chart so you can compare them and what indicator is used for the x-axis. Another option could be that you want to compare different indicators on the same crime, this is easily accomplished as well. Hovering over each bubble tells you what community area you are looking at, and from the chart itself the position and the size of the bubble gives a lot of information already. More on this in section 4.

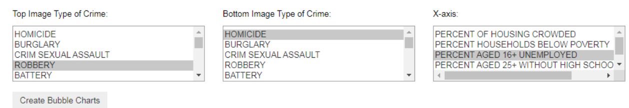

Figure 5: Bubble chart control panel

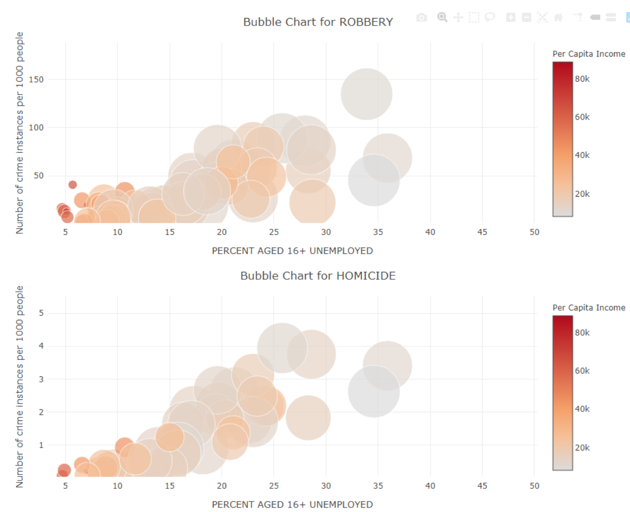

Figure 6: Bubble chart final version

3.4 Interactive Choropleth Map

Given the highly geographic nature of our data we decided to build another tool utilizing a map, this time it is a choropleth map which will help us answer the question 'Where are specific crime types located?'. We have exact data at our disposal about the time, type and location of all reported crimes in Chicago over several years. We combined this with geodata containing the boundaries of the community areas in Chicago to result in an interactive map that will let you see how the hotspots of crime change over the years.

For this choropleth map we use a similar function as applied for the interactive scatter plot but instead of plotting each instance of crime we aggregate based on type and year and color code each community in Chicago depending on the intensity of that particalur crime. We decided to show per community area because they have been static since their inception, while other areas such as districts can change over the years. We then allow the user to interactively decide what year and crime to show through 2 menus shown in figure 7. This type of interactivity lets the user both see how crime locations change throughout the years and see where specific crime types are located.

Once we have aggregated our data to give us the number of crimes per type per community area for each year we can start to visualize it. Here we did not normalize the data relative to the population of

Figure 7: Choropleth map controls

each community area. This was not done because we are interested in finding the main problem areas which is irrespective of population.

Now each area has to be color coded, the first thing we tried was to scale the intensity of the color shown for each community area depending on the crime rate in a linear fashion. However this would mean each community area would have its own color which did not turn out well. Instead we grouped community areas in 5 groups depending on their crime rate. We did this by looking at the highest crime rate and the lowest and then grouping at intervals of 20%. This approach might cause a high outlier to make it seem like there is not much crime because this high outlier would put other not so high crime rates into lower brackets. We however still decided for this approach as opposed to grouping the highest 20% in one group and then the next and so on because that can equally give a skewed view where areas might seem worse then they are. For example if the crime rate is very low in 90% of the areas but high in 10% this approach would not show that well, while the approach currently employed would. Once we have our groups each community area will receive a color depending on what group they are in. These colors were chosen using Colorbrewer to make sure a color mapping is achieved that is perceptually ordered.

An example of a map is shown in figure 8, this map shows the relative robbery rates per community area in 2010. It seems that one community has a relative high crime rate compared to the others, if we are curious whether this is the case for other types of crime as well we can easily find this out. But more on use cases in section 4.

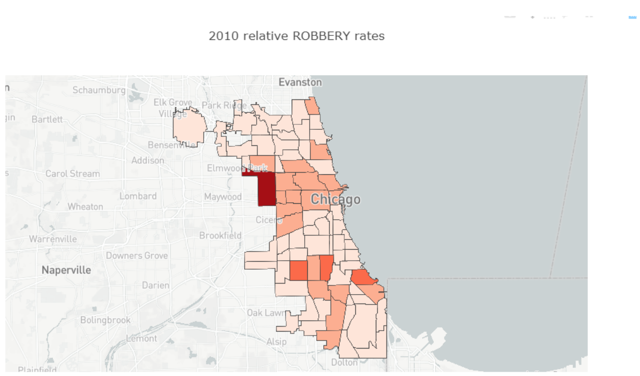

Figure 8: Choropleth map showing the relative robbery rates in 2010 by community area

3.5 Time Based Crime Heat Maps

A heat map is a graphical representation of data where the individual values contained in a matrix are represented as colours. Since we already implemented two tools that focus on answering 'Where are specific crime types located?' we wanted to implement a more information heavy to answer the question 'When do specific crime types occur?'. This is done by creating an interactive heat map that aggregates crimes on times based on location filters. As mentioned in section 3.1 Bokeh has more interactive possibilities, thus this technique was implemented using Bokeh. The movie explorer for IMDB data was used as a framework on which the heat map was based on [5].

The implementation of these visualization techniques shall be split into two parts. First, we shall discuss the interactive widgets and then we shall discuss the heat map.

We chose to make a more dashboard-like overview of the widgets. We adapted, removed and added some widgets from the initial tool. By using the widgets the data can be filtered based on 'Date', 'Arrest', 'Domestic', 'Primary type', 'Location Description', 'Community Area', 'Ward', 'District' and 'Beat'. In addition, we added widgets that change the axes of the heat map, which allows for analysis of crime occurrences based on different time combinations. The options for these widgets are 'hour', 'weekday', 'monthday', 'month' and 'year'. An example would be visualizing a crime heat map for comparing the hour of the day with the month of the year.

Each widget type was carefully chosen to optimize the user experience. Because 'Arrest', 'Domestic' and the axes options only have a few possibilities and because these exclude each other these were chosen to be simple select boxes. The 'Date' was chosen to be implemented using a one-way slider to ensure simplicity of usage. The rest of the filtering widgets were implemented using a multiple selection widget. By doing this we ensure that a user can have combinations of multiple regions, crime and location types, which allows for a broader analysis. Figure 9 shows a configuration using the different type of widgets.

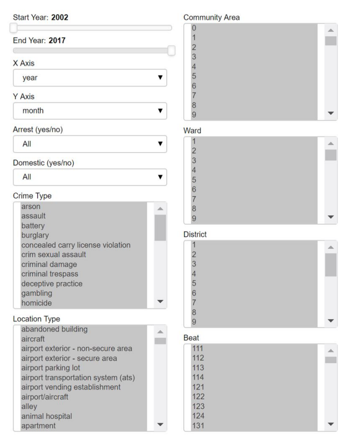

Figure 9: Heat Map Widgets

The actual heat map was constructed using the categorical rectangle function in Bokeh. We decided to remove all complexity of the graph and removed the axes ticks. These ticks were replaced with simple labels stating the category based on one of the main axes labels for time combinations. To ensure readability of the graph we decided to put the x-axis above the figure and let both axes increase from the upper left point. This is done because all information is located in the left or upper part of the graph, and because people tend to read from the upper left corner to the lower right corner according to the F-shape theorem[4].

We also added some more data-centered functionalities. In addition, we chose to show the number of filtered crimes above the graph, because this provides information on the applied filtering and the significance of the plot. We also added a hover function that shows the axes values, count, count deviation from the total mean and count deviation from the axes means. This allows the user to get a better perspective on the data and ensure readability when dealing with plots that have a higher amount of tiles. Lastly, we implemented a colour shading that shows lower values in green colours and higher values in red colours. This shading is accompanied by a colour bar that provides an indication for the values assigned to each colour. This colour mapping, axes and hovers are updated based on the latest filtering applied to the data. Figure 10 shows an implementation of the heat map.

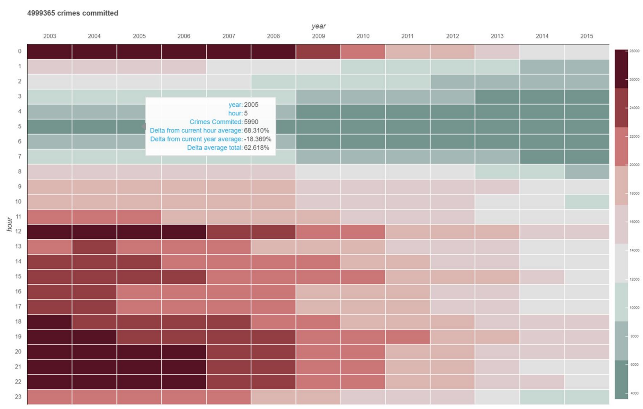

Figure 10: Heat Map

4 USE CASE

Each method described in the visualization design was applied as a visual tool for the analysis of the data set. Due to the complexity and size of the data, we chose to use a part of the data to highlight some interesting findings and show the general potential of each of the visualization methods. In addition, we recorded a screencast of the functionalities of each of these methods to support the claims we make in this paper. The consumer that might be interested in this tool is a chief of police. The next use case findings are thus evaluated as if you wanted to stop crime.

4.1 Interactive Scatter Plot

This section will discuss the findings and use cases for the method obtained in section 3.2. These findings are shown more elaborately in the attached screencast. This visualization method is mostly made for answering the questions 'Where are specific crime types located?' and partly for answering 'When do specific crime types occur?' and 'How do crimes spread throughout the city?' because each crime is shown based on locations after applying a filter.

When not applying enough filters on the data the scatter plot becomes unreadable and messy. So first we decide to filter on a specific year and month. After the application of this filter, we can see crimes more distinctive. When applying another filter on the crimes 'weapon violation', 'prostitution' and 'narcotics' we can clearly distinguish in two north-west and south-west region. You can also see that the crimes originate around these regions and from there spread less dense throughout the city. When hovering over each of these crimes we can then also see additional information about the crime.

When we want to gain insight into the location of the crime we can simply filter on something like 'school'. Then only crimes are shows that occurred on a schoolyard. When zooming in to a specific crime with satellite view you can even see the neighbourhood and buildings surrounding the crime and can thus get a better understanding of the environment of the crime.

Furthermore, we can filter on times, domestic and arrests. Filtering on times also shows a bar/club area in the north-east region of Chicago. Lastly, arrest and domestic crimes also show information about where these happen most. These were some examples of how this visualization tool could be used.

4.2 Bubble chart Animation

This section will discuss the findings and use cases for the method obtained in section 3.3. These findings are shown more elaborately in the attached screencast. This visualization method was made to answer the question 'what causes the different crimes in different areas of the city' and it also partly answers 'where are specific crime types located' but it is not well suited for this. These questions are answered by looking at different socioeconomic indicators as one of the underlying issues causing crime and the amount of crime that happens in each area of the city normalized for population.

We allow for easy comparisons between 2 charts by plotting them next to each other. So when we want to gain insight into for instance the difference between a two very different crime types such as battery and deceptive practice given an indicator we can easily do that as shown in Figure 11. On this image you can immediately see on upward trend on the number of battery cases when poverty rates go up. While apparently deceptive practice has a relative equal number of cases no matter income or poverty levels. And while you could try to explain part of this because the data shows where the crime was committed and not where the perpetrator originated from we could compare deceptive practice with a crime such as burglary that you might expect to be influenced by how rich the area is as well. Yet figure 12 still shows an upwards trend for robbery as well.

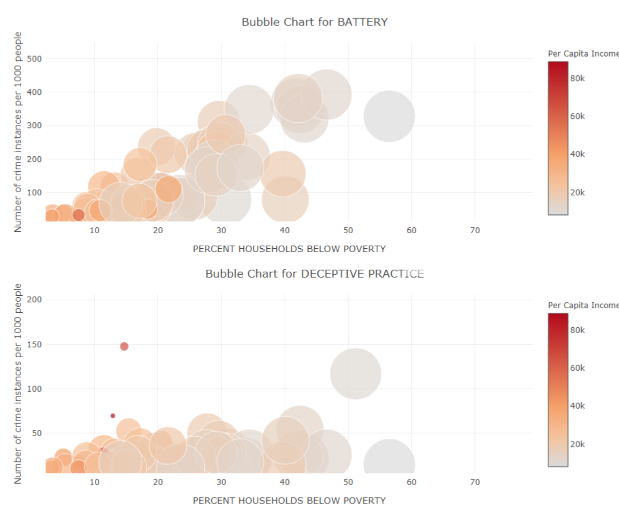

Figure 11: Battery compared to deceptive practice using poverty rates

We can also compare socioeconomic indicators per crime type, this can for instance give us insight into the differences between the effect of poverty and unemployment rates on prostitution. Any many more combinations are possible, more of these can be seen in the attached screencast.

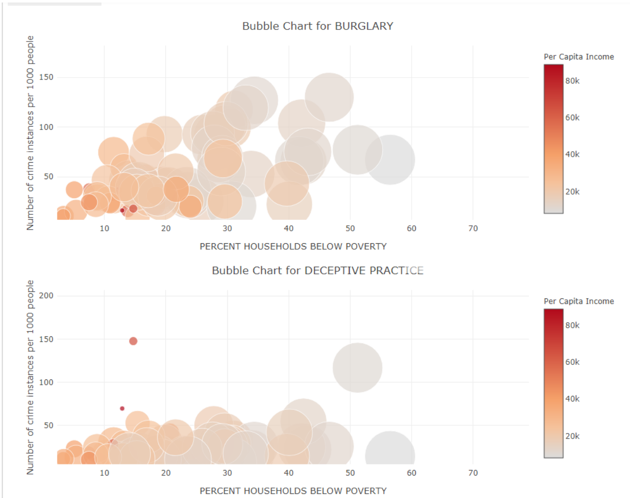

Figure 12: Burglary compared to deceptive practice using poverty rates

4.3 Interactive Choropleth Map

This section will discuss the findings and use cases for the method obtained in section 3.4. These findings are shown more elaborately in the attached screencast. This visualization technique was used to answer the question ”Where are specific crime types located?” but the interactivity allows you to see how the locations of crime changes over the years so it also helps answer ”How do crimes spread throughout the city?”, it can show the spreading of the crime throughout the years. This visualization could also give answers if the user is interested in the change over the years for specific crimes. Or a user might be interested in a specific community area and see how affected by crime that community area is.

To answer ”Where are specific crime types located?” we select a crime type we are interested in and then a year, for instance Figure 13 shows the prostitution rates in 2005, here we can see several hotspots and a decent spread throughout the city. Now there are two directions we can go, we might be interested how the prostitution rates change throughout the years or we might want to see if other types of crime have similar hotspots or if they happen more in other areas. This can allow us to find out if certain crimes are grouped together or not.

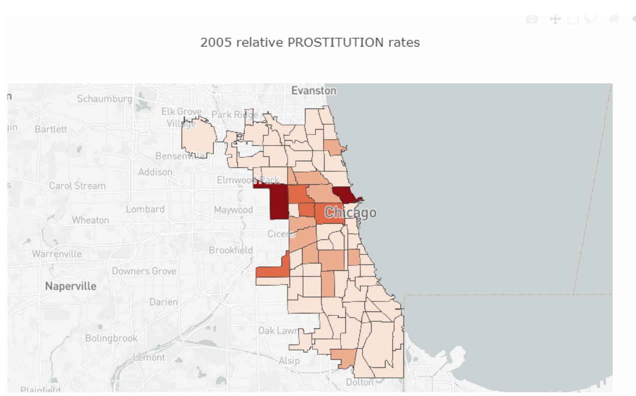

Figure 13: Relative prostitution rates for each community area in 2005

To find out how prostitution changes we can for instance look at Figure 14 that shows the relative prostitution rates in 2015. Comparing this map with the one from 2005 we can see some major changes, evidently prostitution instead of being spread throughout the city has mostly been reduced to only a couple of community areas.

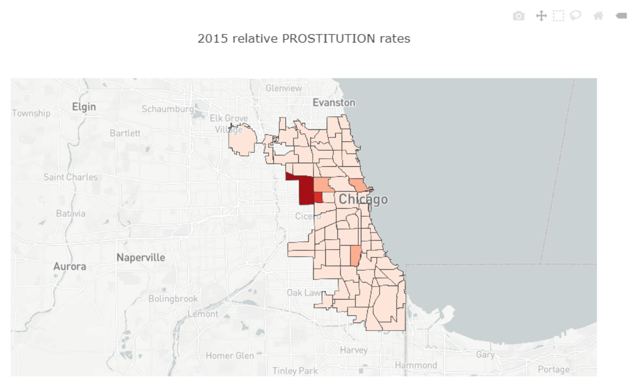

Figure 14: Relative prostitution rates for each community area in 2015

Now we might also be interested in how prostitution compares to for instance human trafficking, you might be interested in seeing if there is a link there. So we create a map for human trafficking in 2015, which is shown in figure 15. It seems there is no direct link location wise, but this only tells us is that the locations of the crimes are different.

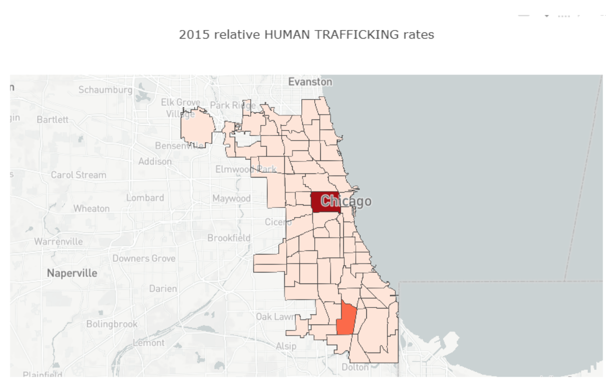

Figure 15: Relative human trafficking rates for each community area in 2015

4.4 Time Based Crime Heat Maps

This section will discuss the findings and use cases for the method obtained in section 3.5. These findings are shown more elaborately in the attached screencast. This visualization method is mostly made for answering the questions ’When do specific crime types occur?’ and partly for answering ’Where are specific crime types located?’ because aggregate crimes on times based on location filters.

Looking at the entire data set 2001-2018 you can see a clear trend of the number of crimes occurring. This is also shown in figure 10. The red colours on the left are replaced with the greener colours on the right of the map. This is thus generally a good thing because the overall amount of crime is decreasing. Sadly, when looking at a specific crime ’theft’ you can see that there has been an overall increase. Because there is a general trend it becomes more difficult to get a clear distinction with 'year' as one of the axes. Therefore we shall continue the analysis on a 'month' and 'hour' level.

When continuing to analyze 'theft' a clear distinction between crimes in summer and winter months can be seen. This is possible due to the hover function, which also shows a percentage compared to the average of the x-axis value. This percentage is increasing in the late hours in winter months. A similar seasonal influence can be derived for gambling in the summer months between 6 pm and 10 pm. Then gambling crimes occur more often around.

When looking at location types some other interesting findings can be derived. Most crimes in bars occur after midnight. In addition, crimes that occur on public transit 'cta' the most occur at 2 pm and 7 pm on weekdays. Interestingly crimes in school-related places occur mostly on weekdays from 8 am and 4 pm, where more crime occur later in the week.

As a Chicago police officer, these insights can all be gained on a level of the 'Community Area', 'Ward', 'District' and 'Beat' for which you or your department is responsible. Using this tool could thus help with aiding the police to station more/fewer employees or have more focus on specific locations or crimes. These were some examples of how this visualization tool could be used.

5 Discussion

The Chicago crime data set has a vast amount of information, due to this there are many directions we could go with this. In the end we settled on four questions we deemed interesting to explore and utilized visualization techniques that we thought would work best to answer these questions. The tools python and jupyter notebooks were chosen because python has a vast amount of visualization libraries at its disposal and jupyter notebooks allows for fast prototyping and visualization.

The interactive scatter plot struggles with the amount of data. This is mostly caused by the fact that mapbox has difficulties with the visualization. In addition, dealing with data as much as a year also causes the plot to overflow with scatter. The scatter plot does offer insight on where and when crimes occur with additional specific data, but it does struggle with the amount of data. Because the data is filtered in the scatter plot we could have also used an interactive density map plot or aggregated map plot, which has also been implemented in section 3.4.

The choropleth map allows us to see where crimes are located through aggregated data. We did not manage to make it work with a legend informing the user about the amount of crime instances which is the major downside of the current implementation. But to answer the question where crime is located it is still well suited. The color coding used even without a legend is understandable due to the order nature of the colors used. The underlying code does struggle with the amount of geojson data it has to process, each community area is described by a vast number of data points to make it as accurate as possible, due to this creating a map usually takes between 5 to 10 seconds which hampers interactivity.

The interactive bubble chart made use of two other data sets to try and give an answer to what causes crimes in certain areas. To do this it aggregated the data from the crime data set and merged this with the other two data sets. It took some time to find data sets that worked with the crime data set. This is also partly the reason why this visualization uses community areas since that is what the other data sets used and we needed a variable to merge on. Now we have quite a bit of data at our disposal, for this we chose a bubble chart since it allows for many variables to be shown in a way that does not occlude the data and presents a nice overview. The decision to show two charts next to each other allows for easy comparison since it is hard for a user to remember exactly what the previous chart looked like. Two versions of the bubble charts were made since the user might be interested in the differences between crimes on the same socioeconomic indicator and in the difference between indicators on the same crime.

The interactive heat map has a good balance between specific insights and aggregated information. The dashboard responded fast and did a good job of dealing with the large data set by aggregation. Beside the aggregation on time periods, we also found out that aggregation-based on the region would make for a good addition. This allows more functionalities when comparing trying to answer questions on where crimes might occur during which times. After implementation we could, for instance, distinguish bar/tavern location crimes per hour of a day per district. This ads a location dimension to the dashboard and result in a good combination of the initially proposed heat map and the density plot map. When combining the heat map with a reference for locations we could even see some pattern about how higher intensity 'narcotics' and 'weapon' crime areas resulted in neighbouring areas also getting higher intensity crime over time.

6 Conclusion

The Bubble chart animation was the only visualization technique that was capable of answering the question 'What causes the different crimes in different areas of the city? and also introduced additional data sets to do so. This framework can be extended with different data sets and allows to simplify the large data set.

The rest of the questions are answered by the other three methods. Here the scatter plot performs well, but is incapable of dealing with the large data set and the complexity of the data. The choropleth map that shows aggregated information-based on a geographic map work better when dealing with the data, but this method loses a bit of diversity and functionality due to the simplification of the data. These two methods did, however, provide more insight into answering 'How do crimes spread throughout the city?'. Lastly, the interactive heat map proposed in section 3.5 with the addition discussed in section 5 proved to be a responsive and useful tool that allowed to answer the question 'Where are specific crime types located?', 'When do specific crime types occur?', 'How do crimes spread throughout the city?'. We propose further fine-tuning and implementation of these heat maps, based on user feedback so that it one day can be used in the actual police department.

7 Breakdown of Work

The work was divided evenly, we worked well together. Each person took upon them to do two visualization approaches which is also detailed in the table below. We kept good communication when we needed help or feedback. Overall it was a nice project to do and we are happy with how it turned out.

David	Mike
Data exploration	Data exploration
Heatmap	Bubble Chart
Scatter Plot Map	Choropleth Map
	Abstract
Section 1	Section 1
Most of section 2	Some of section 2
Section 3,1 3,2 3,5	Section 3,3 3,4
Section 4 4,1 4,4	Section 4,2 4,4
Section 5 and 6	Section 5 and 6
	Section 7
Screencast	Screencast and editing
Final editing	Final editing

REFERENCES

[1] Version 291. URL: `https://data.cityofchicago.org/` (visited on 12/16/2018).

[2] Mike Bostock. *D3.js - Data-Driven Documents*. Version 311. URL: `https://d3js.org/` (visited on 12/16/2018).

[3] *City of Chicago — Data Portal — City of Chicago — Data Portal*. Chicago. Version 294. URL: `https://data.cityofchicago.org/` (visited on 12/16/2018).

[4] *F-Shaped Pattern of Reading on the Web: Misunderstood, But Still Relevant (Even on Mobile)*. Nielsen Norman Group. Version 327. URL: `https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/` (visited on 12/18/2018).

[5] *Gallery fffdfffdfffd Bokeh 1.0.2 documentation*. Version 323. URL: `https://bokeh.pydata.org/en/latest/docs/gallery.html` (visited on 12/18/2018).

[6] *plotly*. Version 307. URL: `https://plot.ly/python/` (visited on 12/16/2018).

[7] *Welcome to Bokeh fffdfffdfffd Bokeh 1.0.2 documentation*. Version 307. URL: `https://bokeh.pydata.org/en/latest/` (visited on 12/16/2018).