

Matriculation Number : 190027923 & 190000096

Overall 982 Words (Excluding reference)

Group Report of CS5044 Practical 3

03.April 2020

Abstract

This report reflects the work of the gun violence incidents visualisation. It first introduces the dataset and some preprocessing steps. Then it explains the visualisation in terms of design ideas, interaction methods and implementation. It also includes possible insights, which our target audience can get from this visualisation. This report ends with a brief reflection and a conclusion.

*** Links to our visualisation :**

https://xs33.host.cs.st-andrews.ac.uk/P3/html_code/index.html

https://cd243.host.cs.st-andrews.ac.uk/P3/html_code/index.html

**** If the links above does not work, please read the "readme" file in the "code" folder, which illustrates how to run our visualisation locally.***

Data Overview and Preprocessing

The visualisation is based on the gun violence incidents dataset, which is downloaded from <https://www.kaggle.com/jameslko/gun-violence-data>, originally from gunviolencearchive.org. Gun violence Archive is a non-profitable corporation, who provides free online public access to information about gun-related violence in the US (GVA). This dataset makes a record of more than 26,000 gun violence incidents in the US between January 2013 and March 2018. This data contains detailed information about each incident, for example location, number of people injured or killed, guns involved, and participants' information.

There are 20 columns in the original dataset. It is not possible to show all features in the visualisation, so we used Tableau Prep to filter out key attributes (Figure 1). Here, each row was copied with different paths for making the Nightingale Rose Chart in Tableau.

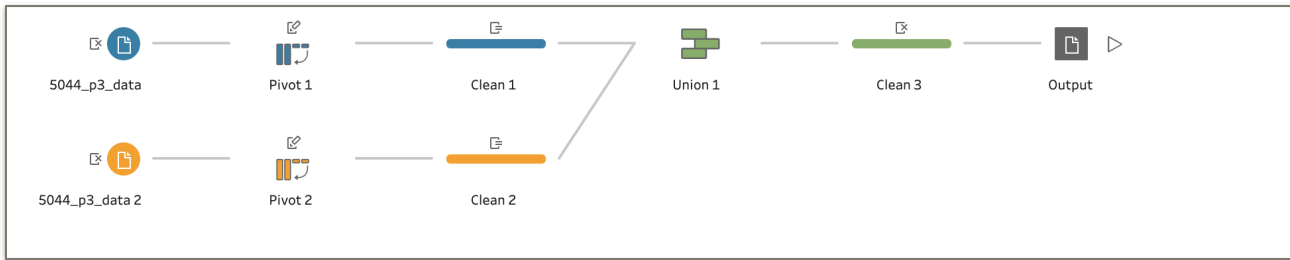


Figure 1: Filtered out key attributes and added paths in Tableau Prep

Then, we imported the data to Python Jupyter Notebook to format the date attribute and sort data by date (Figure 2).

```

data.head()

  Counts  Pivot1 Values  Path    date    state  n_killed  n_injured
0  incident_id      584829  102  19.06.2016   Oregon         1         1
1  incident_id      584966  102  19.06.2016  Nebraska         0         1
2  incident_id      584373  102  19.06.2016 North Carolina         0         1
3  incident_id      583646  102  19.06.2016   Florida         0         0
4  incident_id      583648  102  19.06.2016   Illinois         0         1

data["date"]=pd.to_datetime(data["date"], dayfirst = True)

data.head()

  Counts  Pivot1 Values  Path    date    state  n_killed  n_injured
0  incident_id      584829  102  2016-06-19   Oregon         1         1
1  incident_id      584966  102  2016-06-19  Nebraska         0         1
2  incident_id      584373  102  2016-06-19 North Carolina         0         1
3  incident_id      583646  102  2016-06-19   Florida         0         0
4  incident_id      583648  102  2016-06-19   Illinois         0         1

data["date"].max()

Timestamp( '2018-03-31 00:00:00' )

data = data.sort_values(by="date")
  
```

Figure 2: Formatted and sorted data in Python

Our visualisation is designed to answer the questions: How the Number of Gun Violence Incidents in the US Changes From 2013 to 2018 ? How about the casualties ? Our target audience is the general public, who has little gun violence incidents knowledge. Users are expected to query, search and analyse with the use of our visualisation. The visualisation is based on the below attributes (Figure 3):

ATTRIBUTE	ATTRIBUTE TYPE	DESCRIPTION
Counts	Categorical	incident id name
Pivot1 Values	Categorical	incident id
Path	Quantitative	only for making the Nightingale rose diagram
Date	Quantitative	when the incident happened
State	Categorical	where the incident happened
n_killed	Quantitative	the number of killed people
n_injured	Quantitative	the number of injured people

Figure 3: Key Attributes

Visualisation Description

Design Ideas

Our design strictly followed the expressiveness and effectiveness principle. In other words, the data is visualised without adding ordering information that does not exist and the most important data has been shown via the most effective channel(Munzner, 2014). For example, the number of incidents is a quantitative attribute, so we used colour saturation in the map and the Nightingale rose diagram. Similarly, the number of killed people and the number of injured people are also quantitative attributes. They have different position on a common scale in the line graph. By contrast, the state attribute does not imply a order, so different states have different spatial regions in the map.

Visualisation Interaction

There are mainly two ways to interact with the visualisation. First, the filter fulfills a needed function for the users to chooses from a specific year and the general trend (Figure 4). All diagrams will be updated automatically after users choose a category in the dropdown menu below the main title.

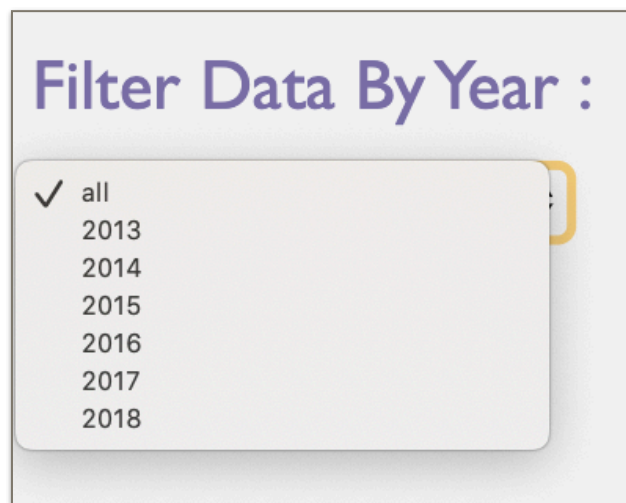


Figure 4: The filter

Second, users can hover over the diagram to get detailed information. For example, users can hover the mouse over sectors in the Nightingale rose diagram to see the ranking information (Figure 5). Users are also allowed to interact with the line gram by hovering over the circles along the paths, then an accurate number will show in the tooltip rectangle. Similarly, when choosing a rectangle in

the heat map, the border of the rectangle will turn into black and detailed information will show in the tooltip top right.(Figure 6).

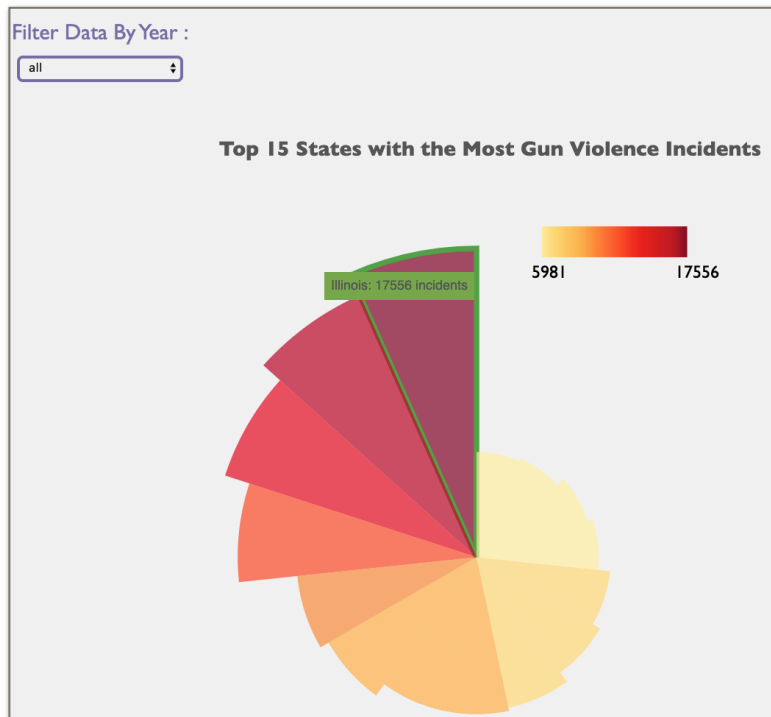


Figure 5 : Hover over a sector to see detailed information

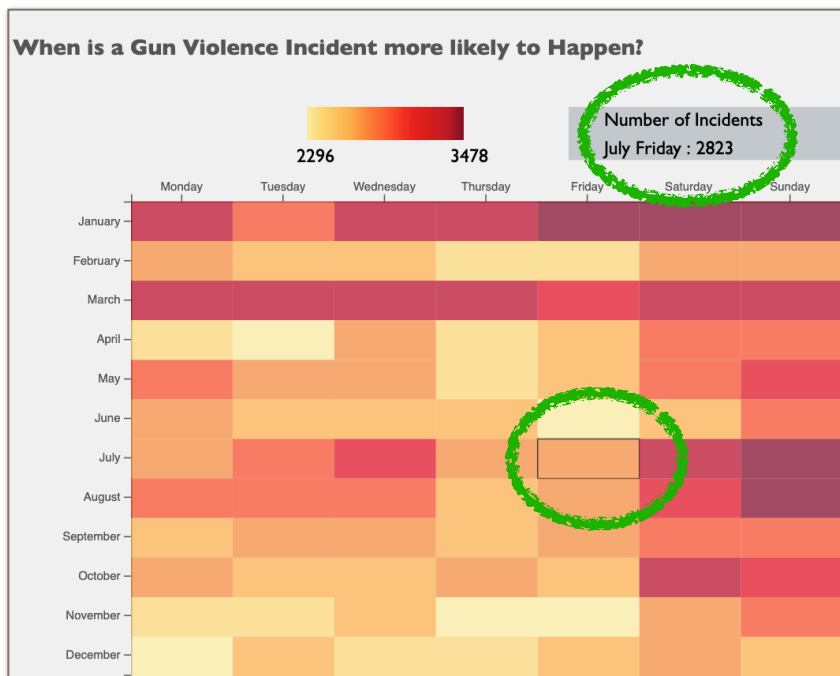


Figure 6: Interact with the heat map

Visualisation Implementation

Our visualisation is written in d3.js. The program first updates the data based on the filter. Then it updates the diagrams. References are listed both below the visualisation and here :

Reference Place	Reference Name	Reference Link
Choropleth Maps Gradient Color	Interactive Choropleth Map	https://leafletjs.com/examples/choropleth/
Filter	week 10 D3 Tutorial dataUpdates	https://studres.cs.st-andrews.ac.uk/CS5044/Tutorials/week_10_D3/D3_tutorial_07_dataUpdates.pdf
Data Processing and Nest	week 8 D3 Tutorial dataTransformation	https://studres.cs.st-andrews.ac.uk/CS5044/Tutorials/week_08_D3/D3_concepts_dataTransformations.pdf
Line Graph	week 8 D3 Tutorial linegraph	https://studres.cs.st-andrews.ac.uk/CS5044/Tutorials/week_08_D3/D3_tutorial_05_linegraph.pdf

Visualisation Insights

There are four diagrams in our visualisation. They focus on three aspects of gun violence incidents. The map and the Nightingale rose diagram are about geographical information. The line gram on bottom left shows casualties trend and the heat map provides time information. The visualisations can effectively answer the questions that we put up.

First, users should choose whether they want to focus on a specific year or the general trend. Then, the map on the top left shows the distribution of gun violence incidents in the US. The change of chromaticity in the map reflects a comparison of the number of incidents in each year. It may be difficult for users to rank in the map, so the Nightingale rose diagram on the right serves as a supplement. Users can also see the relationship between the number of injured and killed people in the line gram. The heat map on the bottom focuses on the most dangerous days. The deeper the colour is on the diagram, the more incidents happened on this day.

Reflection

Our visualisation contains four types of graphs. We try to convey the information more richly. However, there are a few limitations. For example, ways to interact with the visualisation are limited. Adding a filter of states or updating all other diagrams when hovering over the map would make the visualisation more interactive. Moreover, when users run the webpage and operate the

filter function, the loading time is a little bit too long. This may be due to inefficient code execution. Since our team members were working remotely, it is difficult to keep the code style consistent and write code that can be executed efficiently. In order to overcome these difficulties, it would be better to use Git or other version control tools to assist remote programming from the beginning of our project.

Conclusion

To sum up, although there is still room for further improvement, the visualisation successfully introduces the gun violence incidents situation in the US for the general public. In the future, the focus can be switched to adding more interactive elements and improving code efficiency.

Reference

Munzner, T. (2014). Visualisation Analysis and Design. A.K. Peters/CRC Press.