NYCU 2023 Autumn

Data Vislualization
Final Project Report Team23

Bo-Han Chen (陳柏翰)
Student ID: 312551074
bhchen312551074.cs12@nycu.edu.tw

Xu Lin (林煦)
Student ID: 312553027
f94061042.cs12@nycu.edu.tw

Abstract

In this project, we develop a data vislualization system for raffic accident data in Taiwan in 2022. Our goal is to find out several insights related to the cause and trend of traffic accident, which can help the experts and government to improve traffic safety. With our system, the user can easily find out the temporal and spatial trend of the accident, additionally, the driver-contributed cause and vehicle information are also included, which can be used for further analysis.

Motivation

Since there are many people injured or killed in traffic accidents every year, so the analysis of the traffic accident data is important and necessary to prevent the accident. Some of the accidents are caused by the driver's behavior, and some are caused by the road condition and the circumstance such as weather condition at the time that the accident happened. Therefore, an easy-to-understand data vislualization can help the experts to analyze the cause of the accident and promote the right policy to drivers and pedestrians. By analyzing the circumstance of the accident, we can also find the dangerous road condition and time period, which can help the government to decrease the accident rate by improving the road condition and employing more traffic police at the certain time period.

Questions to Answer

The following questions are what we want to answer by visualizing the dataset.

1. Which time period has the highest accident rate?
2. Which weather condition mainly causes the accident?
3. Is there any relationship between the driver-contributed causes and the type of vehicle?
4. Which part of the vehicle is most likely to be hit and fragile during the crash?

Dataset

The dataset we choose is the Taiwan road traffic accident statistics in 2022 (*111 年傷亡道路交通事故資料*), which contains the details of the accidents that cause death or injury in Taiwan. The dataset is public and can be accessed from Open Data Website of Taiwan Government [1].

Data Description

The dataset contains 845547 records, including 4544 A1 level records and 841003 A2 records. The A1 level means the accident causes death in within 24 hours, and the A2 level means the accident causes injury or death in more than 24 hours. Each record contains 51 attributes, including the time, location, weather, road type, vehicle type, and the number of death/injury. Some details such as whether the driver is drunk or not, which part of the vehicle hit in the accident, and the driver-contributed cause of the accident are also included.

[1] `https://data.gov.tw/dataset/161199`.

Methodology

Figure 1 shows the overview of our vislualization system. The system is devided into 6 parts, including Taiwan traffic accident map, line chart of death / injury trend, stream chart of death / injury trend on different weather condition, heatmap of accident rate, stacked bar chart of driver-contributed cause and vehicle type, and the crash position distribution of vehicle.

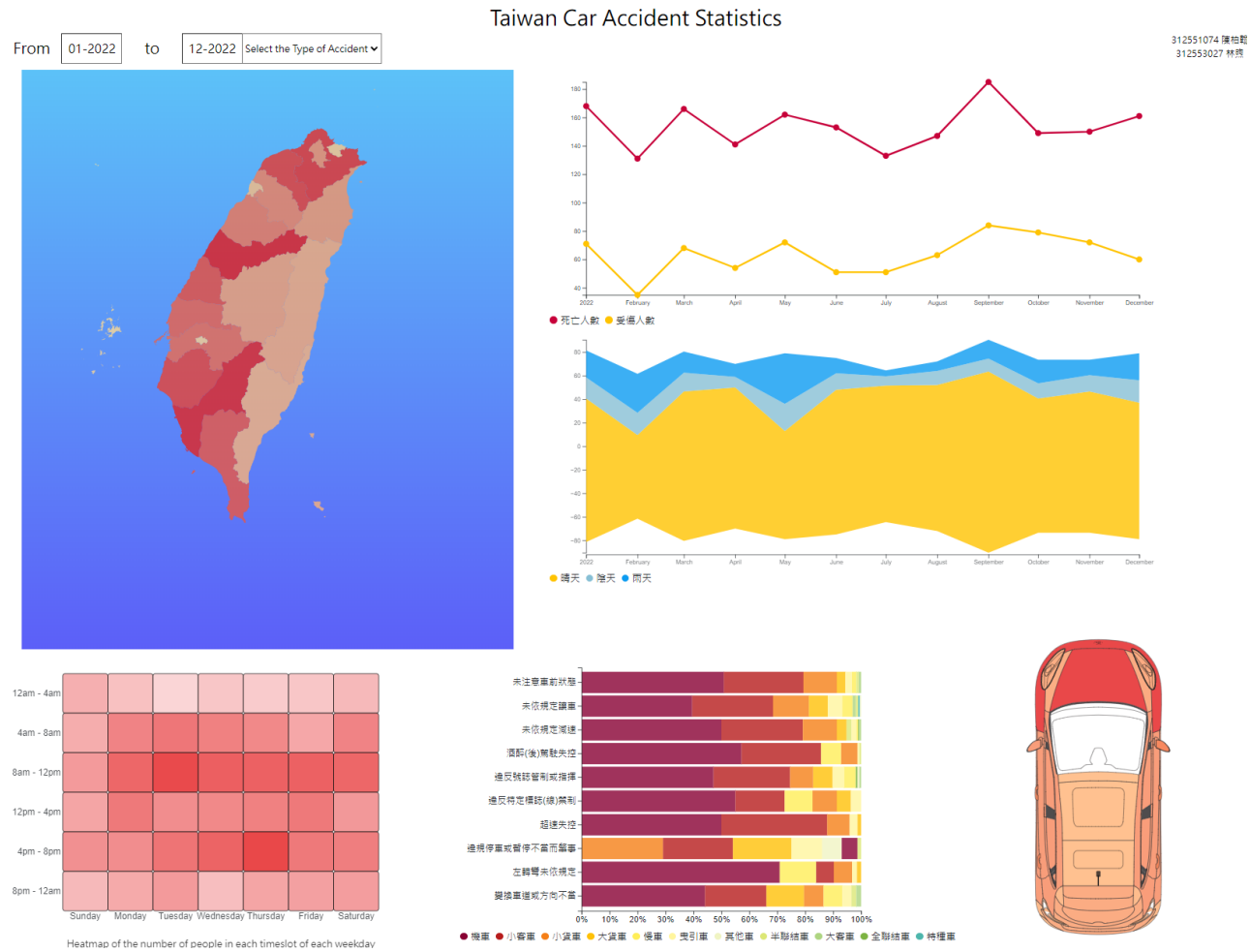


Figure 1: System Overview

Taiwan Traffic Accident Map

This map shows the death / injury count of each cities in Taiwan. The user can directly indicate which city has high fatality counts in the selected time period with the ordinal color scheme. For detailed information, the user can move the mouse to the city and the tooltip will show the death / injury count, drunk driver count, and the mobile phone usage count, which shows in Figure 2.

Line Chart of Death / Injury Trend

This line chart shows the death / injury trend in the selected time period, which can help the user to find out the temporal trend of the accident. The chart will also show fine-grained information when the selected time period is small, For example, if the user selects the time period of 1 month, the chart will show the trend of each day in the selected month, if the selected period is larger than 1 month, the chart will show the weekly trend. The tooltip shows the time period and death / injury count of each data point, which shows in Figure 3.

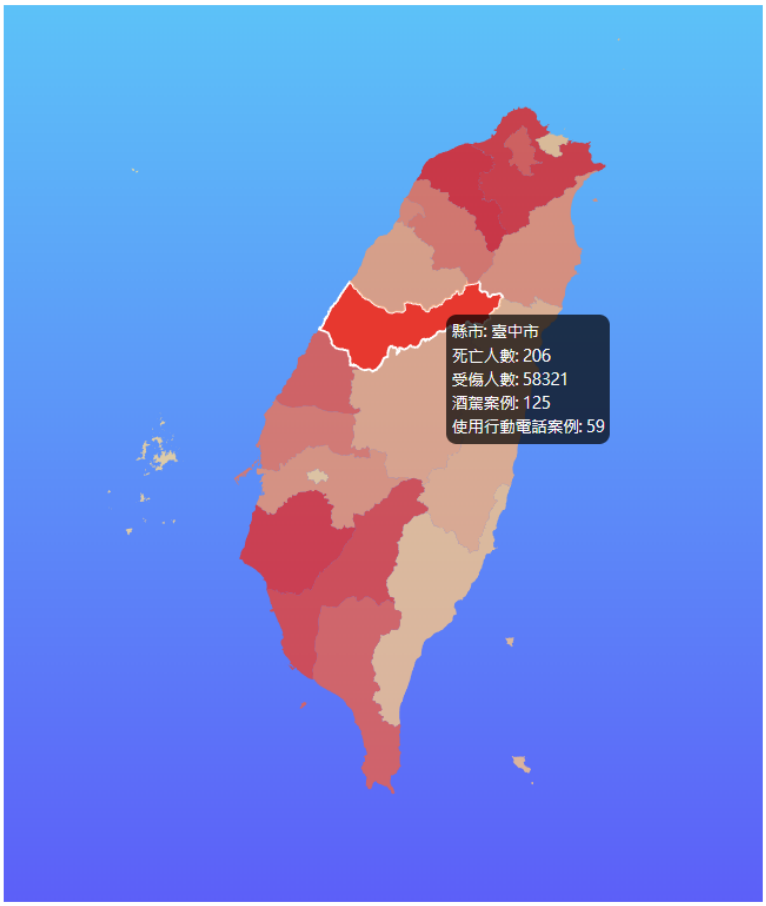


Figure 2: Map Tooltip

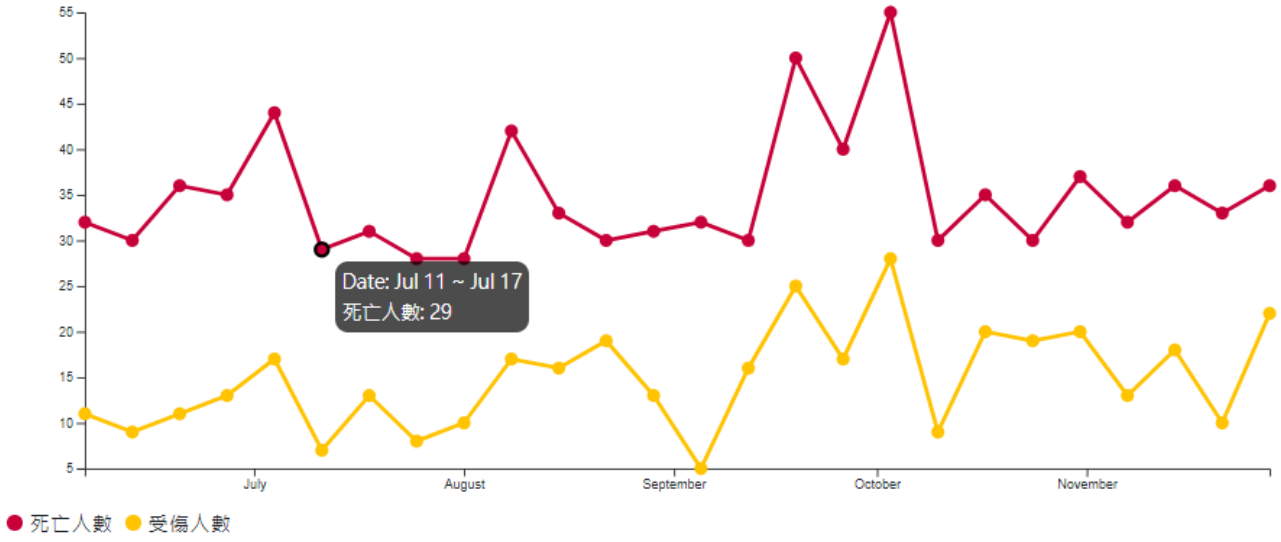


Figure 3: Line Chart Tooltip

Stream Chart of Accident Count on Different Weather Condition

This stream chart shows the accident count on different weather condition. User can find out the accident trend on different weather condition. The weather condition is divided into 3 categories, including sunny, cloudy, and rainy. Figure 4 shows the stream chart in our system, the tooltip shows the time period and accident count on different weather condition.

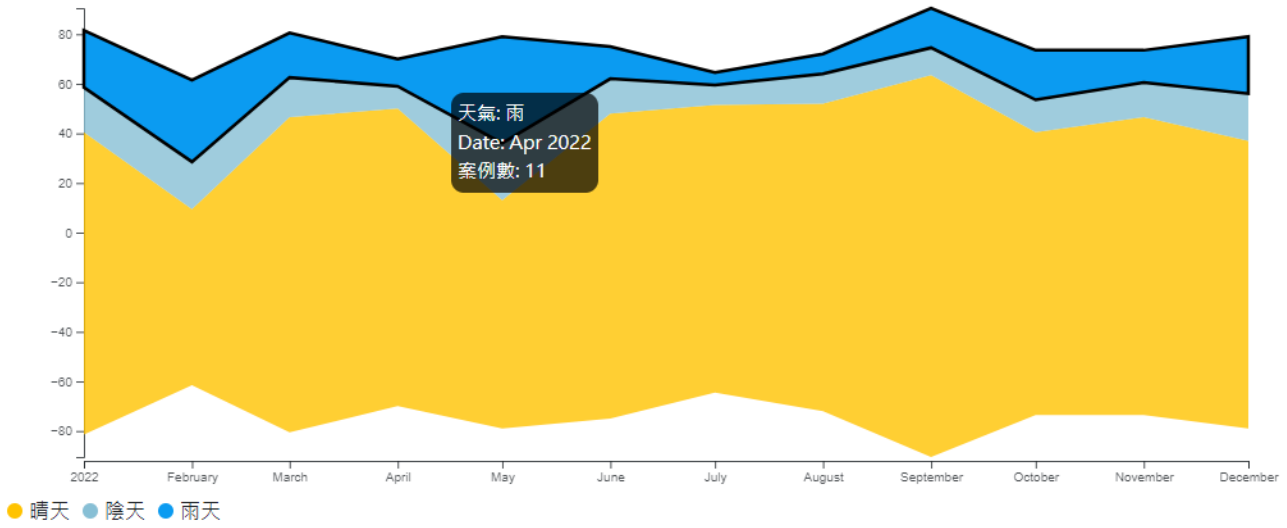


Figure 4: Stream Chart Tooltip

Grid Heatmap of Accident Count

The heatmap shows the accident count by time of day and day of week. This helps user to find out the cyclical trend of the accident, such as comparing the accident count between weekday and weekend, and the accident count in the morning and afternoon. Figure 5 shows the heatmap in our system, the tooltip shows the time period and accident count of each cell.

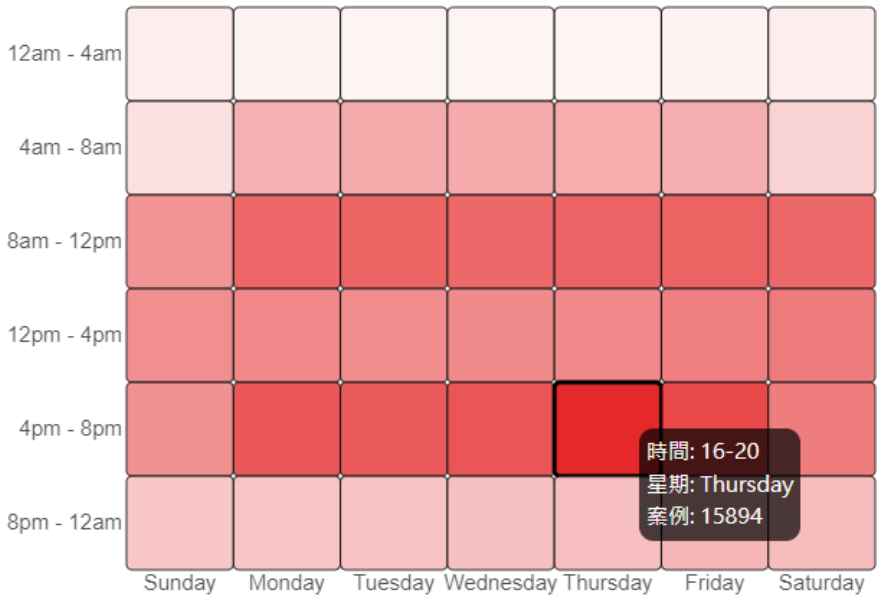


Figure 5: Heatmap Tooltip

Stacked Bar Chart of Driver-Contributed Cause and Vehicle Type

This stacked bar chart shows the ratio of vehicle type for each driver-contributed cause. User can first find out the frequent driver-contributed cause, which is sorted by the total accident count. Then,

user can find out the vehicle type that is most likely to be involved in the corresponding cause. Figure 6 shows the stacked bar chart in our system, on this chart, we can find out that the most frequent cause is ” 未注意車前狀態”, and the tooltip shows that scooter is the most likely vehicle type to be involved in this cause, which is 50.9% of the total accident count.

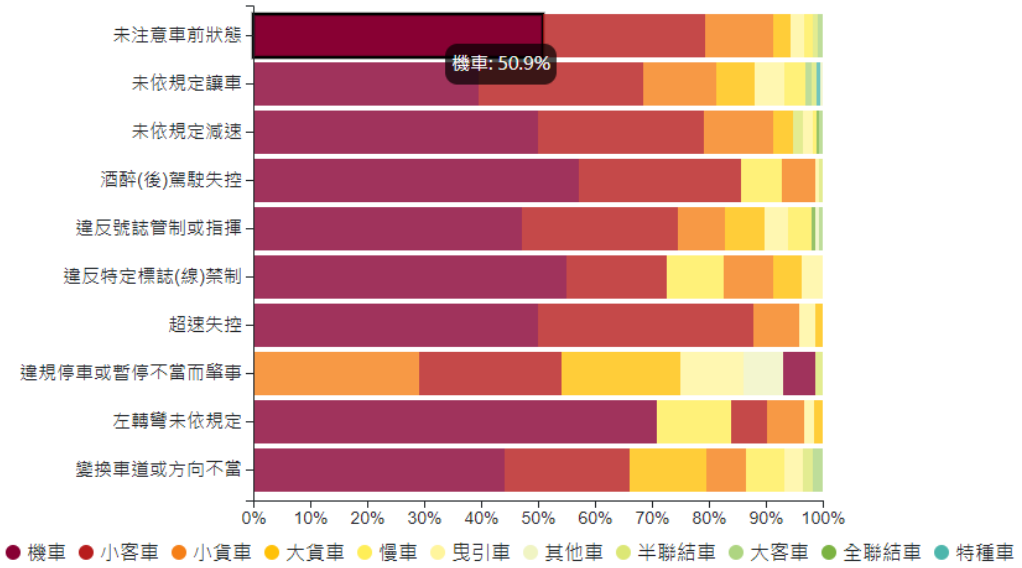


Figure 6: Stacked Bar Chart Tooltip

Crash Position Distribution of Vehicle

This chart shows the crash position distribution of vehicle. In our dataset, the crash position is divided into 8 categories, including front, rear, left, right, front-left, front-right, rear-left, and rear-right. User can find out which part of the vehicle is most likely to be hit and fragile during the crash by the ordinal color scheme. Figure 7 shows the crash position distribution of vehicle in our system, the tooltip shows the crash position and the accident count.

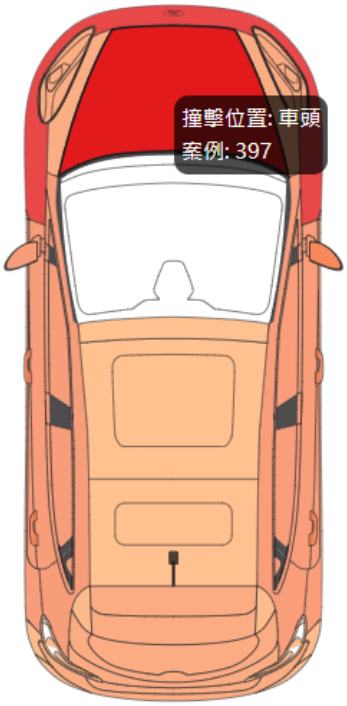


Figure 7: Crash Position Distribution of Vehicle Tooltip

Interaction

The user can select the time period and type of accident data (A1, A2, or both) in the control panel on top of the system (Figure 8). Furthermore, city selection is also available in the map, by clicking the city, system will show the corresponding information.

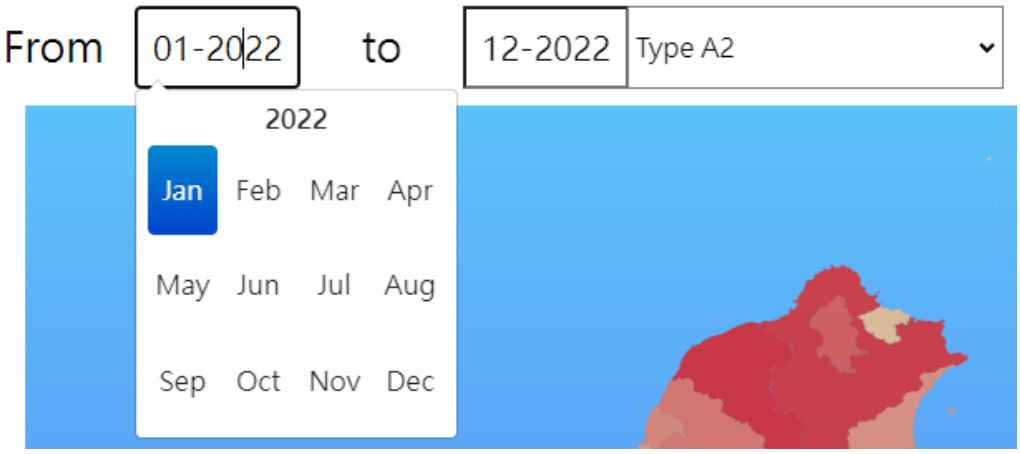


Figure 8: Control Panel

Insights

Temporal & Spatial Trend

By selecting both A1 and A2 data, we can first find out the city that has the high fatality and injury count in 2022 is New Taipei City, TaoYuan City, Taichung City, and Tainan City in Figure 9. each of them has more that 58000 injury count and 160 death count. Comparing the casualty count between different cities in Taiwan, we can find out that the count in eastern Taiwan is much lower than the other cities, which we can infer is because the traffic density is lower in eastern Taiwan.

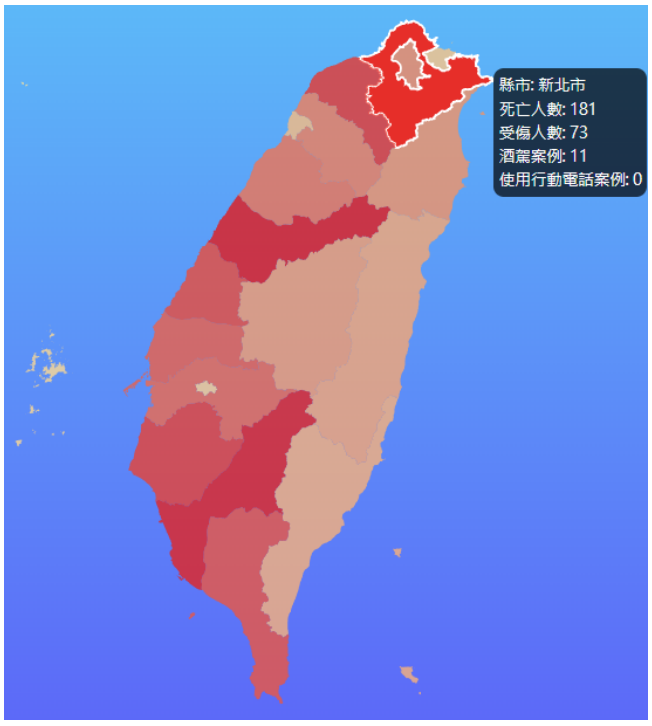
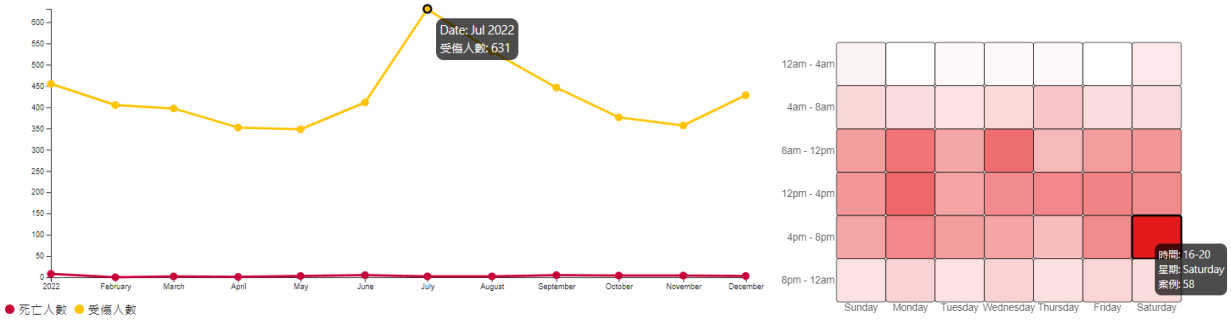


Figure 9: Accident Trend in Different City

We can further find out the accident trend in eastern / western Taiwan. By selecting eastern cities such as Taitung, we can find out that the peak of the accident count is in july and August from Figure

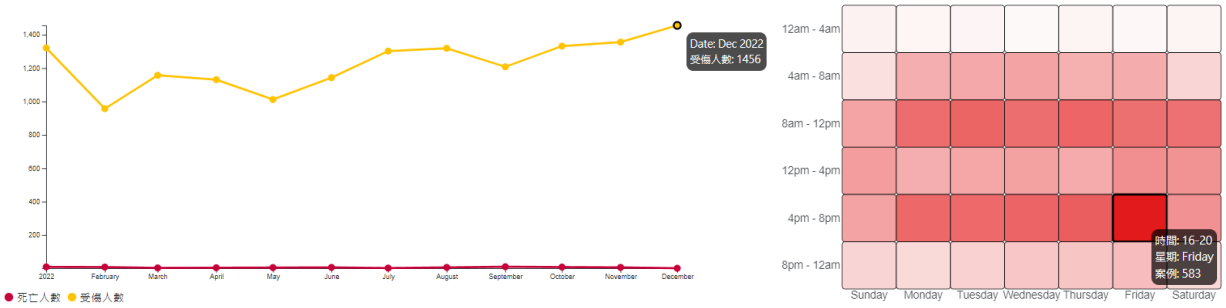
10a, and the accident count on weekend is abnormaly high shown in Figure 10b. We infer that the reason is there are many tourists in eastern Taiwan during summer vacation, and the traffic density will increase during the weekend, which causes more accident than other time period.



(a) Accident Trend in Taitung (b) Accident Count in Taitung

Figure 10: Accident Trend in Taitung

For western cities such as Hsinchu, the trend can be found in Figure 11. We can find out that the casualty count becomes higher in winter from Figure 11a, this may cause by the early sunset and rainy weather in winter. From Figure 11b, we can notice that the accident count is higher in the 8:00-12:00 and 16:00-20:00, which is the time period that people go to work and go home, so the traffic density is higher than other time period, which causes more accident. Furthermore, the accident count on Friday in period of 16:00-20:00 is higher than other weekday, we think the reason is that people may go to other city for travel or go back to their hometown during the weekend, so the traffic density is higher than other weekday.



(a) Accident Trend in Hsinchu (b) Accident Count in Hsinchu

Figure 11: Accident Trend in Hsinchu