NYCU 2023 Autumn

Data Vislualization
Final Project Proposal Team23

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

Xu Lin (林煦)
Student ID: 312553027

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 vechile hit in the accident, and the driver-contributed cause of the accident are also included.

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 and light 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 such as road and light condition, we can also find the dangerous road section and the 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. What is the most dangerous road type?
2. Which time period has the highest accident rate?
3. Which circumstance mainly causes the accident?
4. Is there any relationship between the driver-contributed causes and the type of vechile?
5. Which part of the vechile is most likely to be hit and fragile during the crash?

Methodology

The overview of our visualization system is shown in Figure 1.

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

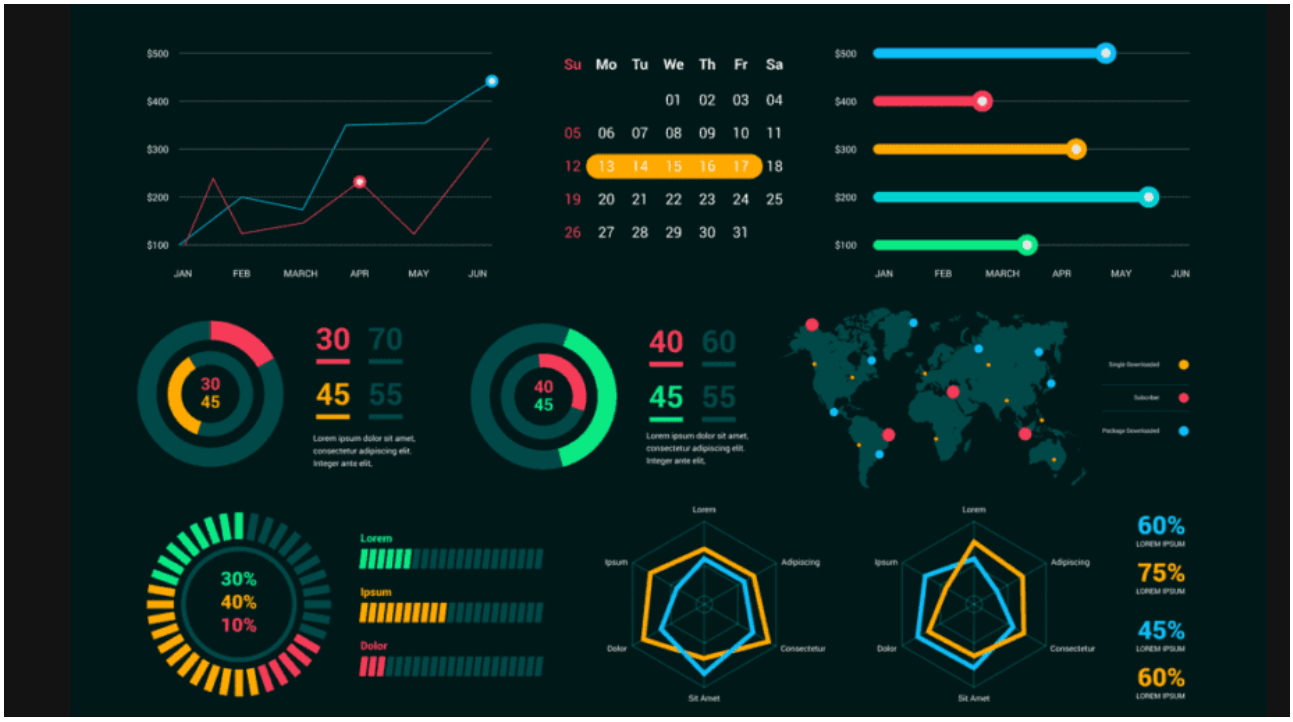

Figure 1: Overview of our project

Taiwan Map

The map shows the number of accident in each city by heatmap, with darker color means higher accident rate. The user can move the mouse to the city to see the statistics details such as the ratio of drunk driving and mobile phone usage.

Accident Statistics Over Time

The line chart is used to let user discover the trend of the accident occurence over months and days. And user can also select and focus on specific time period. However, it's not enough to depicts the trend of accidents from the perspective of months and years. So we also use the grid heatmap to represent the accident occurence by time of day and day of week. We expect it can provide another detailed view of traffic accident trend.

Vechicle and Driver Information

The stacked bar chart is used to show the relationship between the driver-contributed causes and the type of vechile. With this chart, we can find the most frequent cause of the accident and the type of vechile that corresponds to the cause.
The car chart shows the distribution of crash positions of the vechile. With this chart, we can find the most frequent crash position of the vechile and the crash part that causes more death with A1 level data.