Northeastern University

IE6600: Computation and Visualization for Analytics

Project 2: Traffic Accident Analysis of Tempe

Group 17

Ashna Arora

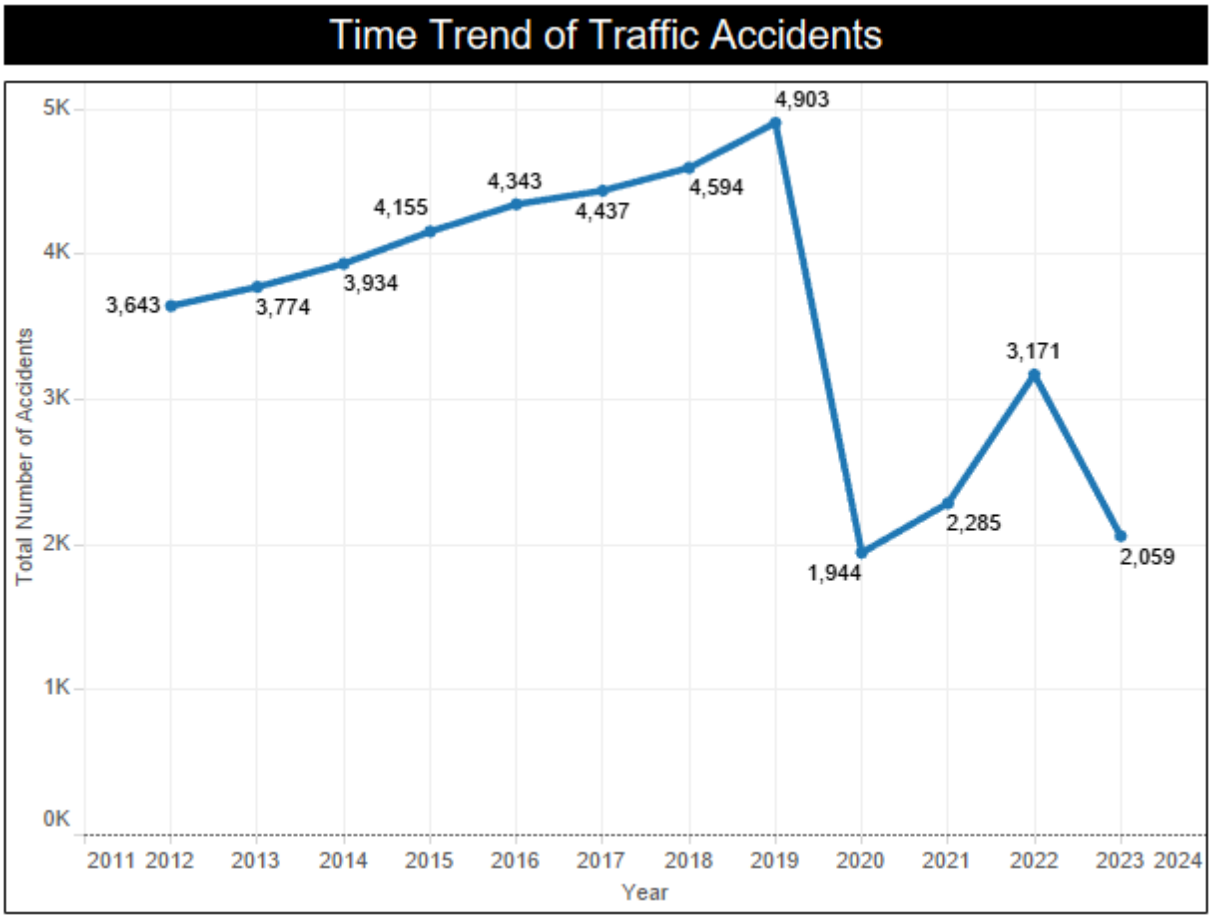
Dheeraj Kumar Goli

Siyuan Gao

1. Introduction

Traffic accidents are a major public health issue globally, leading to significant loss of life, injuries, and economic costs. Understanding the factors contributing to road accidents is crucial for developing effective prevention strategies. This report delves into the "Crash Data" dataset from Tempe, a comprehensive collection of 43,242 traffic accident records, through exploring the influence of various factors on traffic accidents, this report puts forward some suggestions to reduce the traffic accident rate effectively.

2. Problem Statement



From the time trend chart, we can find that the traffic accident rate was relatively high before 2020, and the traffic accident rate decreased after 2020 because of the novel coronavirus. In order to prevent the traffic accident rate from returning to the previous high level, what specific measures can we take to reduce the traffic accident rate? What are the external and internal factors that cause traffic accidents?

3. Data Description

3.1 Data Source:

Dataset from data.gov

3.2 Data Description

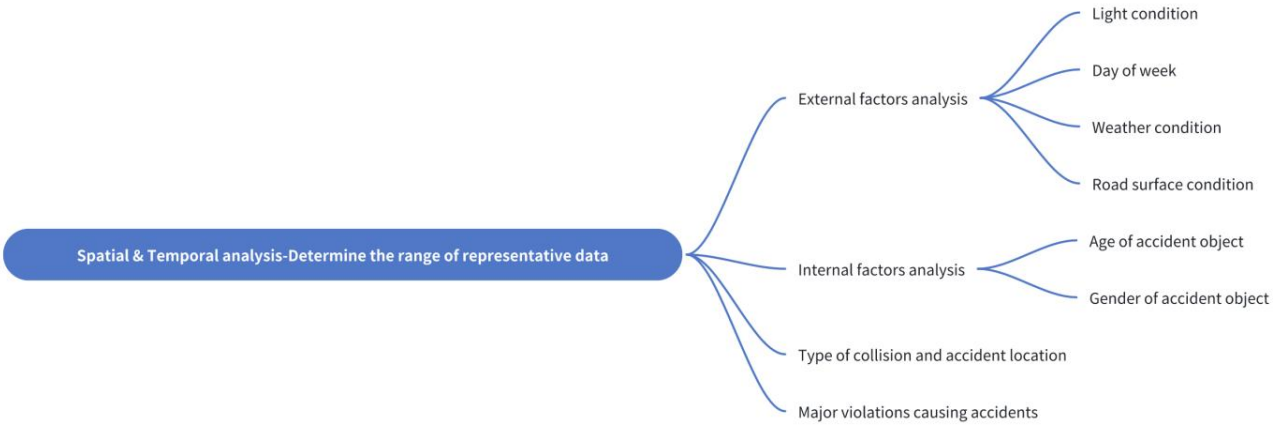
Column Name	Interpretation
OBJECTID	Accident ID
DateTime	Time of Accident
Year	Year of Accident
StreetName	Street Name of Accident
CrossStreet	Cross Street Name of Accident
Totalinjuries	Total Number of Injuries
Totalfatalities	Total Number of Deaths
Injuryseverity	Accident Injury Degree
Lightcondition	Light Condition of Accident
Weather	Weather Condition of Accident
SurfaceCondition	Surface Condition of Accident
Unittype_One	Type of Accident Object 1

Age_Drv1	Age of Accident Object 1
Gender_Drv1	Gender of Accident Object 1
Violation1_Drv1	Interpretation of Action of Accident Object 1
AlcoholUse_Drv1	Alcohol Use of Accident Object 1
DrugUse_Drv1	Drug use of Accident Object 1
Unittype_Two	Type of Accident Object 2
Age_Drv2	Age of Accident Object 2
Gender_Drv2	Gender of Accident Object 2
Violation1_Drv2	Interpretation of Action of Accident Object 2
Latitude	Latitude of Accident
Longitude	Longitude of Accident
hour	Hour of Accident
day_of_week	Day of the Week of Accident

3.3 Data Processing:

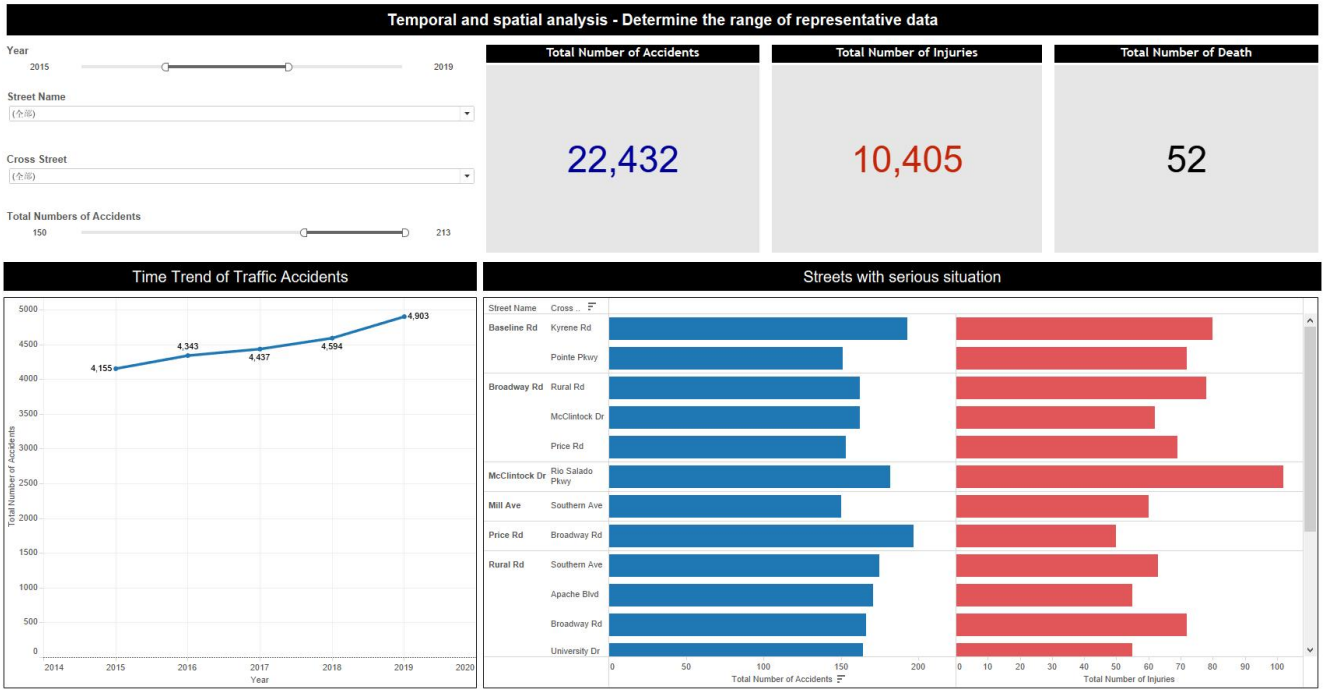
We selected the key columns and dropped the null value. To ensure the integrity of the data, we filtered the outliers in the Tableau platform using the filter function.

4. Methodology and Design Process



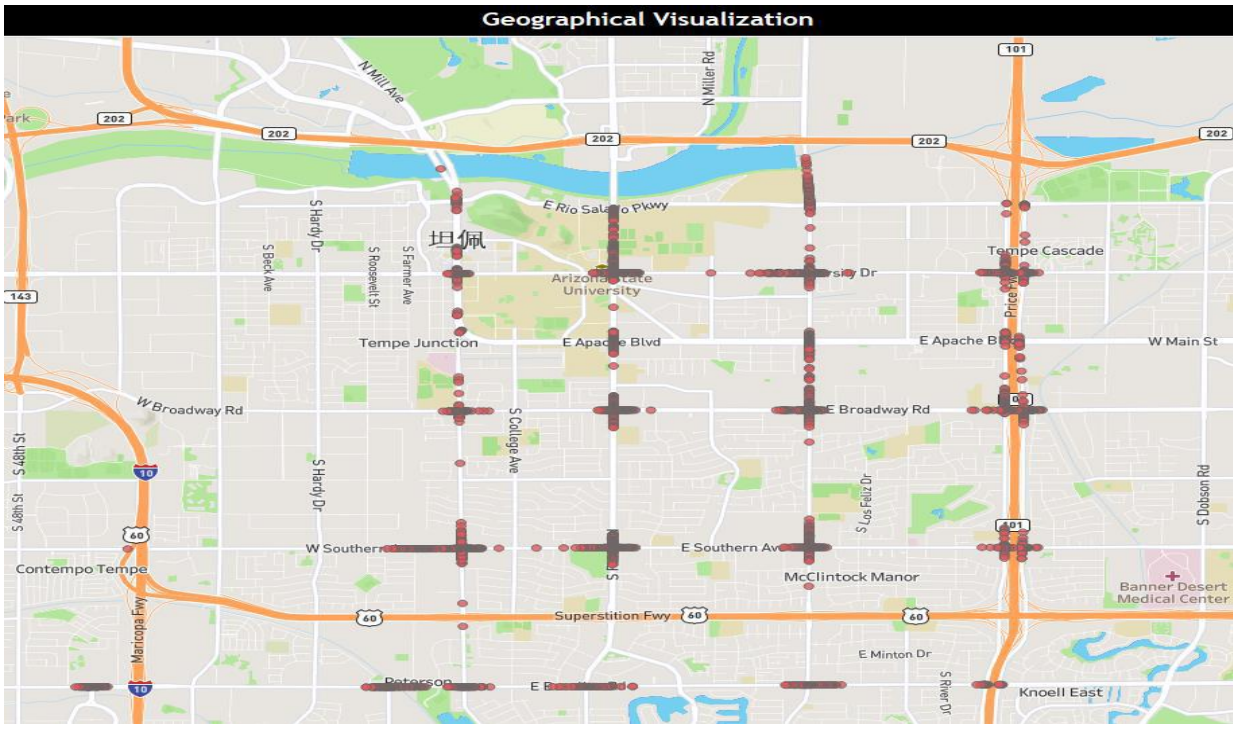
4.1 Determine the range of representative data

Before 2020, the number of traffic accidents is on the rise year by year, and 2015-2019 is the five years with the largest number of traffic accidents, so we choose the data of these five years as the basis for subsequent analysis. Then to identify key locations, we filtered the number of accidents down to at least 150, and we got the eight worst streets, and all the streets that intersect them.



4.2 Geographical Visualization

Here we present all the filtered data in the real map, so that you can have a more intuitive sense of the location and frequency of traffic accidents.

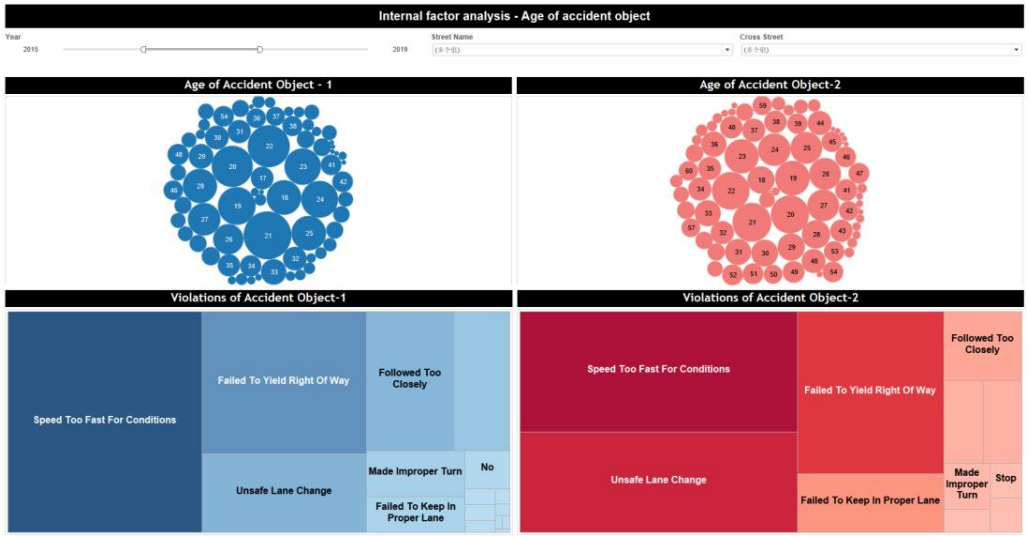


4.3 External factors analysis

In the external factor analysis dashboard, we can see that the traffic accident rate increases significantly between 7pm and 4am, which means that the traffic accident rate is higher under poor lighting conditions, and there is a higher accident rate on weekdays than on weekends, mainly because there is a higher travel demand on weekdays. In addition, we can see that the traffic accident rate is not significantly related to weather and road conditions, and most traffic accidents occur on clear and dry roads.

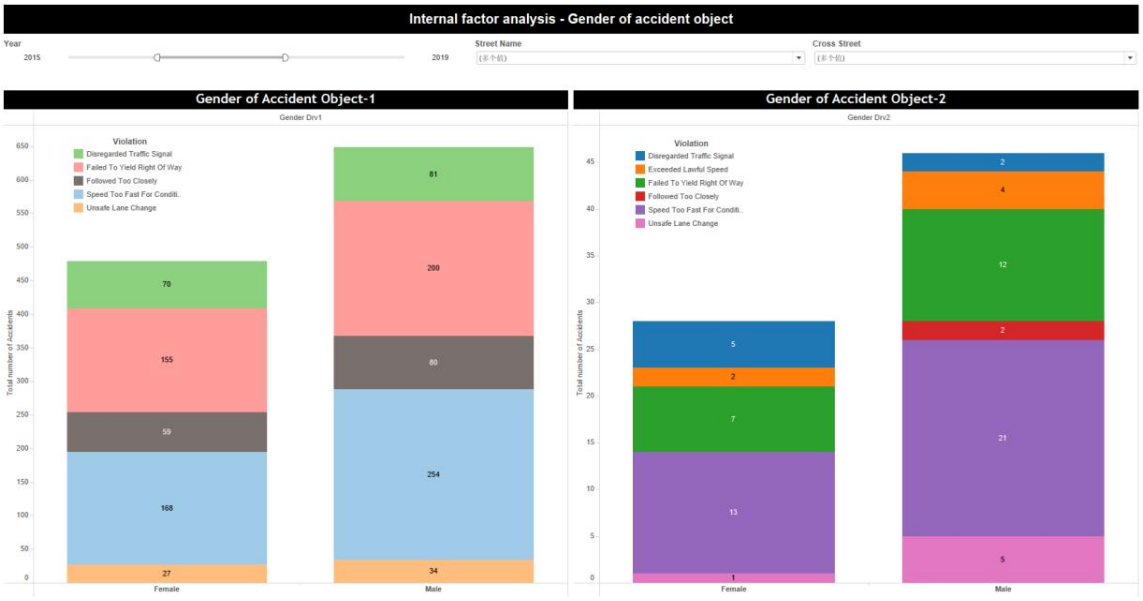
4.4 Internal factors analysis-Age

In the age analysis of the accident objects, we can see that most of the accidents are caused by young people between 18 and 25 years old, mainly because of speeding and refusing to yield the right of way. Most of the passive victims of accidents are also young people aged 20-28, mainly due to speeding and unsafe lane changes.



4.5 Internal factors analysis-Gender

In the gender analysis of the accident objects, we can see that the traffic accident rate of men is about 33% higher than that of women. No matter what kind of violation, the accident rate of men is higher than that of women, indicating that male drivers tend to have more impulsive driving behaviors.



4.6 Type of collision and accident location

In the collision type analysis table, we can see that most traffic accidents are vehicles, and in the following location analysis thermal map, we can get a specific location distribution. For example, the intersection of University Dr and Rural Rd is the most dangerous, with 213 traffic accidents occurring in 5 years.

4.7 Major violations causing accidents

From this word cloud map, we can see that speeding accounts for the largest proportion of all accident types, followed by refusal to yield the right of way, and unsafe lane changes. If the subsequent traffic management is to be carried out, then these three reasons are the worthiest of attention.

5. Key Insights

- There are 18 intersections with the highest traffic accident rates and most noteworthy:

Street Name	Cross Street Name	Total Accidents	Total Injuries
Baseline Rd	Kyrene Rd	193	80
	Pointe Pkwy	151	72
Broadway Rd	Rural Rd	162	78
	McClintock Dr	162	62
	Price Rd	153	69
McClintock Dr	Rio Salado Pkwy	182	102
Mill Ave	Southern Ave	150	60
Price Rd	Broadway Rd	197	50
Rural Rd	Southern Ave	175	63
	Apache Blvd	171	55
	Broadway Rd	166	72
	University Dr	164	55
	6th St	159	81
	Rio Salado Pkwy	150	69
Southern Ave	Rural Rd	194	85
	Mill Ave	159	65
University Dr	Rural Rd	213	78
	McClintock Dr	167	54

- The traffic accident rate is higher under poor lighting conditions, and there is a higher accident rate on weekdays than on weekends.
- Most of the accidents are caused by young people between 18 and 25 years old, mainly violations are speeding and refusing to yield the right of way. Most of the passive victims of accidents are also young people aged 20-28, mainly due to speeding and unsafe lane changes.
- The traffic accident rate of men is about 33% higher than that of women.
- For car-car collisions, we should pay more attention to 4 intersections:
    - Price Rd - Broadway Rd
    - Baseline Rd - Kyrene Rd
    - Southern Ave - Rural Rd
    - University Dr - Rural Rd
- Speeding too fast and refusing to yield the right of way are the most frequent violations in traffic accidents.