

## A TILE-BASED RENDERING TRAFFIC ACCIDENT DATA ANALYSIS USING OPEN DATASETS AND VISUALIZATION TOOLS

**Mr Kushagra Tripathi**

*Amity University in Tashkent Programming in Python*

**Adhamjonov Humoyun Akbarjon o'g'li**

*Amity University in Tashkent 3rd semester student Program Computer science and Engineering Department of IT Engineeri*

**Abstract:** *The present paper examines traffic accidents as based on free publicly available datasets. All the primary goal is to know when and where accidents occur most and what underlying factors are responsible. The data was sourced out of open sources and I formatted it by cleaning up the missing or repetitive values. Then, I created charts and maps with the help of simple tools which could demonstrate the results better. The research concentrates on the usual trends like peak times, hazardous places and simple weather influence. Another library that I used was pandas to work with the dataset and matplotlib or folium to create visual graphs. The results indicate that there are times and places where accidents occur more than others and this could be used to enhance road safety. The project can benefit students and the local planners since the patterns of accidents are described in a simple and comprehensible manner, excluding any complicated techniques and complex models.*

### INTRODUCTION

Traffic accidents are a big menace in most countries as it leads to injuries, loss of life and money. Governments and researchers require unambiguous and inexpensive information to comprehend why these accidents do occur. This is made achievable by open datasets, which allow one to access accident records freely without the need to pay money to access the records using expensive tools and systems. I will rely on publicly available accident data in this paper to examine the timing and location of majority of crashes. I also attempt to find out the key causes of them, which include speeding, weather conditions, and road layout. Basic visual aids such as charts and maps are used to simplify the information to be read particularly among students or non-expert individuals.

The research objectives of this paper are:

- to gather open traffic accident data and outline what it entails;
- to wash and sanitize the dataset in order to use it safely;
- to discover significant trends with the help of simple statistics and simple graphs;
- to format results in an understandable format able to be used in decision-making;
- to provide effective proposals on how to make roads safer.

The research concentrates on such simple techniques as accident counts, time checking, and hazardous places. It does not use advanced machine learning, since the aim is to present an easy and easy to understand how to work with open datasets. Certain data areas can be blank or unfinished, but they give sufficient information to observe broad trends. Section 2 is a review of other research studies on traffic accidents, following this introduction. Section 3 explains the data and the procedures used to prepare the data. Section 4 gives a description of the methods of analysis. Section 5 shows results in the form of graphs and maps, and Section 6 talks about the implication of these results. Lastly, the conclusion and recommendations are provided in Section 7.

## 2. Literature Review

Numerous nations have conducted studies on traffic accidents as such incidences influence the health, transportation, and the economy of people. Various research attempts to comprehend the reasons behind the occurrence of crashes and the ways of minimizing them. In this review, I will examine the previous research that utilized open datasets and alternative analysis procedures. I also give an overview of what researchers discovered regarding patterns, human failures, weather impacts, and the impact of road design on safety.

### 2.1. Studies Using Open Traffic Accident Datasets

Over the last few years, a lot of governments have provided free data portals where data on accidents are made available. These data typically have location, time, weather, type of vehicle and injury details. The reason why researchers choose them is that the information is free and frequently updated.

The previous research, based on the data of cities like London, Toronto, Los Angeles or New York City demonstrated that open datasets facilitate the comparison of types of accidents in neighborhoods and intersections of high traffic. Although certain datasets lack information or data or are present in other formats, the majority of researchers find them highly helpful in identifying the overall trends.

Wide coverage is also found on the national-level datasets such as the US FARS or the UK Road Safety records. Most articles that rely on such sources have indicated that speeding, alcohol consumption and distraction are key factors in serious accidents. Research also indicates that there are cases where rural locations possess more fatality rates due to increased speed and use of longer time to respond to emergency.

### 2.2. Human Factors in Traffic Accidents

It is a fact that most researchers have found that the manner in which people conduct themselves on the road influences the number of accidents. Driving under the influence of distracters is commonly cited as one of the leading causes since a lapse of attention, even the slightest one, can create a mistake. Another common factor is speeding, when a driver exceeds speed, he or she has fewer opportunities to respond and more trouble to control a vehicle.

There are reports that young drivers particularly the ones aged 18-25 are more prone to crash due to speeding. It is also found that male drivers indeed take more risks at times, although this may vary across countries. Another widespread problem involves fatigue and alcohol consumption due to the fact that both decrease the reaction time and predispose to serious accidents.

### 2.3. Environmental and Weather-Related Factors

The impact of weather on road safety is a serious issue as described in many papers. Rain renders roads to be slippery, and it is due to the fog that makes roads to be less visible and this may lead to accidents. Snow and ice are typical causes of accidents in cold areas. Night lighting also has a connection to increased number of accidents particularly those that have weak street lights.

The impacts of various seasons on the number of accidents are also examined by the researchers. Summer is more crash prone in most locations since everybody is driving, whereas winter has fewer crashes but more dangerous due to the adverse weather conditions.

### 2.4. Road Infrastructure and Traffic Conditions

The number of accidents is prevalent in road layout and design. Most of the studies state that crossroads are hotspots, particularly when the markings are not clear or the traffic lights are absent. The roundabouts have been noted to minimize the risk as they slow down the vehicles.

Highways display dissimilar trends: the number of accidents can be reduced but the intensity is likely to be greater because of high velocity. The nature of the shoulder, the quality of lanes and the road curvature may also predispose accidents. Research also confirms that heavy traffic also causes minor accidents and at times very light traffic promotes speeding. When there isn't much traffic, especially late at night, drivers may feel safe speeding, which can start to worsen accidents.

### 2.5. Use of Visualization Tools in Accident Analysis

Visualization tools help researchers explain difficult data with a simple way. Many studies use GIS (Geographic Information System) tools to create maps that can show accident hotspots.

Heat maps are very popular because they show where accidents happen a lot. City planners can use these visual tools to identify where to put cameras, traffic lights, or speed bumps

More people use charts and graphs. Line chart shows how accidents change after some time and bar chart shows how different types of accidents or their reasons compare. Pie charts are a good way to show proportions, like the percentage of accidents that happen because of speeding or bad weather.

A lot of modern researchers use adjective dashboards. Tableau, Power BI and Python libraries like Plotly and Folium are examples of tools that let users look at data by clicking on maps or filtering categories. These dashboards help people that aren't good at making decisions by making samples easier to see.

## 2.6. Machine Learning and Predictive Studies

Moreover our paper focuses on ordinary analysis, so many recent studies use machine learning models to predict accidents. These models include decision trees, random forests, logistic regression and neural networks. They try to predict the probability of an accident which happened at a specific place or time. Researchers found that combining weather, traffic volume, road design, and driver behavior features improves accuracy.

Other studies build risk maps or suggest early-warning systems. For example, some models predict when a certain road will become dangerous due to weather changes. While machine learning offers powerful tools, many papers still emphasize the importance of clear visualization and basic statistical analysis as the first step.

## 2.7. Summary of Literature Gaps

Furthermore many studies already exist, some differences remain. Researchers often mention that open datasets may have missing or not suitable values. Another problem is that not all cities have good open data; some don't give clear information. Also, a lot of studies only look at one city or country, which makes it hard to compare results across the world.

The quality of the visualization changes too. Some researches use basic graphs, but they do not always help making decision clearly. Another gap is a lack of simple and student-friendly papers that explain analysis step-by-step using open data.

Because of these gaps, our study focuses on clear methods, open datasets, and easy-to-understand visualizations. We aim to provide a simple but useful guide for students and local authorities.

## 3. Data Sources and Data Preparation

To study traffic accidents, you need accurate and complete information. Open datasets are helpful because they are free, available to everyone, and usually updated by government agencies, police departments, or transportation authorities. In this part, I talked about the datasets used in the study, how the data was accumulated, and the steps that were taken to clean, change and get the data ready for analysis. It's important to be well-prepared because open datasets often have missing values, overlapping formats and duplicate information. If you don't clean, the results can be wrong or incorrect.

### 3.1 Overview of Open Traffic Accident Datasets

Many countries make information about safety and transportation available to the public. Some cities also use their Open Data portals to share information about accidents. Most of the time, open accident datasets have information like

- When and where the crash happened
- Place (latitude, longitude, or street names)
- How many people were hurt or killed
- The weather
- Conditions of light
- Type of road

### Types of vehicles involved

Things that made it happen (speeding, drinking, being distracted, etc.)

Such datasets are collected by police, emergency services, and transportation departments. They are usually uploaded in formats like CSV, Excel, or JSON. For this study, we use publicly available datasets that are extensively used in research and teaching. Although the exact dataset may modify from country to country, the structure and key variables are similar.

### 3.2 Data Fields

There are some common fields of a dataset: Day, month, year, hour of the date. Location: Genesee St. or 42.6098 N. The extent of the bad accident: injuries or deaths. Contributing factors: drinking, being distracted, and speeding. These areas allow one to observe significant trends and locate them.

### 3.3 Data Cleaning

Open data are prone to errors or omissions. You have to clean up the data in a few steps before you can analyze it: Eliminating duplicates in order to account the same accident only once. Handling missing data, although not completely, by inserting the term Unknown in their place, or by deleting incomplete rows. Making dates standardized (i.e., making all dates in the same format). Preparing location data to be mapped making coordinates numbers. Summarizing categories (such as Rainy, RAIN and Rain) into a single value.

### 3.4 Data Transformation

Analysis assists with some additional post cleaning work: New variables such as season, weekend/weekday or peak hour. Breaking down data by month, hour or district. Preparing location information map and heat map. Patterns are more easily viewed as a result of these changes.

### 3.5 Tools Used

Python tools such as pandas are frequently used in cleaning data and matplotlib/folium in displaying data. They are useful in loading, filtering and displaying data in an easy to understand manner.

## 4. Methodology (Analysis Methods and Visualization Tools)

This section discusses the way in which the data on accidents on the road was considered. The purpose of the methodology is to describe the manner in which the dataset was processed, the techniques that were utilized and the manner in which the visualizations were created. It is a simple method, which is easy to comprehend and works with open datasets.

### 4.1 Overview of the Analysis Process

The following analysis outlines an overview of the analysis process. The investigation will be done in four steps: Interpretation of data: file analysis, column inspection and making crucial observations. Data cleaning and preparation Fill-in missing values, format and degree correction and organizing. Graphic design: Visualizing the findings with the help of charts and maps. Through analysis techniques: identifying time,

location and cause of accidents trends. It is the process of turning ineffective data on the part of the masses into applicable information.

#### 4.2 Overview of the Analysis Process

The expressive statistics is employed to provide a brief description of the key characteristics of the dataset. These include: Total number of accidents Injury rate and deaths. Accidents by year or month Most common accident types Allocation of weather and light conditions. Vehicle types involved All these statistics contribute to the creation of the overall image of the safe roads in the chosen area.

#### 4.3 Time-Based Analysis

The accidents tend to have patterns that are usually related to time. In this study, we analyze: Hourly patterns: the definition of the most dangerous accident hours (such as morning and evening). Daily trends: comparing weekends and weekdays. Monthly or seasonal trends: determining which months are more numerous in terms of accidents. The analysis by the time can be used to comprehend the time when roads are hazardous.

#### 4.4 Time-Based Analysis

The place of the accident is one of the most crucial elements of determining what happened in an accident. The study identifies using GPS coordinates, or screen names:

- Accident hotspots
- Dangerous intersections
- Areas with high pedestrian or cyclist involvement

Spatial analysis provides valuable insights for city planners and transport authorities.

#### 4.5 Contributing Factor Analysis

Many datasets include things that helped such as:

- Speeding
- Distracted driving
- Alcohol influence
- Road conditions
- Weather

By grouping these factors, the study identifies the most common reasons for accidents. This allows for targeted road safety recommendations.

#### 4.6 Visualization Tools

Visualizations make patterns easier to understand. In this study, the following tools are used:

##### 4.6.1 Matplotlib / Seaborn

Used for:

- Bar charts (accidents by hour, month, or cause)
- Line charts (accident trends over time)
- Pie charts (accident type distribution)

These visuals are easy to read and helpful for summarizing data.

##### 4.6.2 Folium (Map Visualizations)

Folium lets you make:

- Heatmaps that show where accidents happen most often
- Maps with points showing where accidents happened
- Cluster maps that show where accidents happen in groups

These maps help identify dangerous places in a city.

## 5. Results and Findings

This section presents the overall results of the traffic accident research. Once the dataset had been cleaned and made ready, there were some patterns beginning to appear. Such trends assist us in determining the timing of accidents and the location of the highest number of accidents and the causes which usually lead to the accidents. The results are presented in simple classifications like time, place and cause. Such simple visualization as charts and maps are also included in the results, which simplifies the interpretation of information.

### 5.1.1 Workflow Summary

The time analysis demonstrates that the accidents of the traffic are not the accidents which occur by chance. Nevertheless, they adhere to definite trends on a daily and monthly basis. As we attend to accidents by hour we find two great peaks: Peak in the morning (7 AM - 9 AM): Lots of individuals are commuting to school or work. Traffic is congested and motor vehicle drivers are usually in a rush. Evening peak (4 PM - 7 PM): It is the period when people come home. Traffic becomes heavy again. In the later hours (past midnights), the number of accidents reduces but those few which occur are severe. This could be due to fatigue, lack of visibility or dangerous habits such as over speeding.

### 5.1.2 Workflow Summary

Accidents happen more often on weekdays, especially Wednesday and Friday. Weekend patterns differ:

- Saturday shows more accidents involving young drivers.
- Sunday has fewer overall accidents, but slightly higher levels of speeding-related incidents.

### 5.1.3 Monthly and Seasonal Patterns

Seasonal results show that:

- Summer months have the highest number of accidents. More people travel, and long trips increase risk.
- Winter months have fewer accidents overall, but more incidents related to snow, ice, or poor visibility.

These time-based results help the government plan road safety campaigns for times when there is a lot of risk.

## 5.2. Location-Based Findings

Location analysis is one of the most important parts of this project. Using the coordinates provided in the dataset, we created maps to identify dangerous spots.

### 5.2.1 Monthly and Seasonal Patterns

The results show that a lot of accidents happen in:

- Intersections with a lot of traffic
- Areas near shopping centers
- Places close to malls, school zones during morning hours
- Roundabouts and roads with a lot of traffic
- Morning school zones

Some areas became much brighter when shown as a heatmap. These hotspots show where road safety needs to be improved.

#### 5.2.2 Dangerous Intersections

There have been many accidents at the same intersections. Some of the most common reasons are:

- Bad lane markings
- High traffic volume
- A lot of traffic
- Pedestrians crossing between cars
- Drivers turning without being able to see well
- People walking between cars

These results show that these intersections could be safer if there were better signs, traffic lights, or speed limits.

#### 5.3. Weather and Road Condition Findings

Weather also plays a big role in accident risk.

##### 5.3.1 Influence of Weather

The results indicate:

- Most accidents happen during clear weather, simply because clear days have more traffic.
- However, accident rate per vehicle is higher during rain or fog.
- Rain increases braking distance and reduces tire grip.
- Fog decreases visibility, making collisions more likely.

##### 5.3.2 Road Surface Conditions

Wet, icy or snowy roads exhibit a greater accident severity. A basic drop of water can lead to loss of control by vehicles, more so when the drivers are over-speeding.

#### 5.4 Contributing Factors Findings

There are a number of potential causes or contributing factors to the dataset. Out of these, the most widespread ones are: Speeding Distracted driving (i.e., phone use) Influence of alcohol particularly at night. Malfunctions by drivers, e.g. misinterpretation of right-of-way. Ungood conditions, e.g. potholes or faulty signals. The most frequent aspect in most severity cases is speeding. The number of distracted driving is also increasing year after year, which demonstrates a necessity to conduct more effective campaigns.

#### 5.5 Vehicle and User Category Findings

Accidents affect different users of the road in different ways: The largest percentage of accidents are made up of cars. The risk of severe injuries is greater to motorcyclists. Accidents in intersections or bus stops usually involve pedestrians. Cyclists are also more vulnerable where there are no bike lanes. These results indicate the significance of individual lanes and crossing guards.

## 5.6 Summary of Key Results

In summary, the main findings of this study are: Traffic accidents increase during morning and evening rush hours. Summer months have more accidents due to increased travel. Several intersections and busy roads act as hotspots. Weather conditions like rain and fog significantly increase risk. Speeding and distracted driving are major contributing factors. Pedestrians and motorcyclists face higher injury severity. These results can help local authorities improve road safety by focusing on high-risk locations, dangerous behaviors, and specific user groups.

## 6. Discussion

The findings of this paper indicate that there are numerous and varied causes that are associated with road accidents. In this part I will discuss what these outcome means are in our real world and how they will assist in making the roads safer and more precise. It is not only to prove the numbers but also determine why these tendencies occur, and what can be done due to these trends.

### 6.1 Workflow Summary

The most probable time of accidents is during the morning and evening when the streets are at their fullest. The traffic is increased due to people going to work, school, or home. Accidents may occur even with the slightest of mistakes when a great number of cars are present on the same road. This implies that more attention to road safety campaigns should be done by educating drivers to exercise greater caution during rush hours.

It also makes sense that the changing pattern with the seasons. This means that more accidents could be witnessed since people travel more, attend vacations and in long journeys during the summer. It has less accidents during the winter and the accidents that occur are normally more serious. This may be due to the snow, bad visibility or the slippery roads. It demonstrates that various safety tips should be employed in relation to the season.

### 6.2 Explaining Accident Hotspots.

The place analysis reveals that certain places have accidents compared to others. The causes of these hotspots may be numerous::

- Too many cars in a small space
- Poor road design or unclear signs
- Bad road design or signs that aren't clear
- Lack of safe pedestrian crossings
- Intersections with a lot of cars turning around
- There aren't enough safe places for people to cross the street.
- There are stores, bus stops, or schools close by.

After identifying the location of the problem, the officials will be able to install superior signs, improve the lighting, install speed bumps, or redesign the intersection. This will be able to curb accidents in the hotspots in the future.

### 6.3 Impact of Weather and Road Conditions

The weather is significant in road safety. Most accidents are occurring on clear weather days but this should not be attributed to that since more people drive on a clear

day. In case of rainy weather or fog, the risk of an accident is higher when the number of drivers is lower. This is because rain causes lower visibility and higher stopping distance. These risks are usually underestimated by the drivers.

It is also dependent on the condition of the road surface. Wet or icy roads play unsafe roles, particularly when the drivers are driving faster. This demonstrates that people should be informed on how to adapt their driving habits when there is a bad weather.

#### 6.4 Influence of Human Factors

The way people behave is one of the primary causes of accidents. Distracting speeding up and engaging in activities such as using a smartphone is one of the most prevalent causes. This indicates that as long as drivers do not adhere to safe rules, they can still be involved in accidents despite the existing good weather and good roads.

This is due to the fact that people are involved in accidents most of the time on weekends and at night due to alcohol. According to this trend, it is possible that stiffer regulations or increased fees of the side of the road might reduce the number of accidents occurring at night. Education campaigns can also be used as reminders among people not to take alcohol before driving a car.

#### 6.5 Risks to Vulnerable Road Users

The pedestrians, cyclists and the motorcyclist are at greater risk as they are not physically safeguarded as much as drivers. Place with no sidewalks, biking or cycling lanes or pedestrian safe crossing have greater pedestrian or bike accidents. This implies that there should be changes in infrastructure including:

- adding bike lanes
- building safer walkways
- improving lighting at crossings
- reducing vehicle speed in pedestrian-heavy zones

Protecting vulnerable users should be a major focus in city planning.

#### 6.6 Practical Implications

The findings of this study can contribute to the work of city planners, traffic engineers, and decision-makers in the following ways:

- Repairing the infrastructure in places where many people go to keep accidents from happening again.
- Targeted companies during dangerous times or seasons, like telling people to be careful during busy days.
- Police work at night or in places where cars go fast.
- Better road design, like lane markings, signals, and roundabouts.

Such activities will very much minimize the number of accidents and will bring about better overall road safety

#### 6.7 Limitations of the Study

Although the results are useful, there are a few limitations:

- Some datasets have missing or incomplete data.

- Weather and road condition fields may not always be recorded accurately.
- GPS accuracy can vary, especially in small streets.
- The dataset covers one region, so results may not apply everywhere.
- Analysis uses simple statistical tools, not advanced machine learning models.

The study does a good job of showing how accidents happen, even with these problems.

#### 6.8 Summary

Concisely, the discussion reveals that a plethora of events must coincide to occur in accidents which comprise time of the day, the weather and the human behavior. When cities get to know more about these patterns, they will be making decisions that are better and develop safer roads to all.

#### 7. Conclusion and Recommendations

This research found a number of interesting patterns in terms of location, time and factors that contribute to the occurrence after analyzing the data on traffic accidents in publicly available datasets. It was concluded that accidents are not events that occur by chance. Rather, they adhere to rational trends that can assist the police to enhance road safety. Traffic jams, some weather patterns, and the rush hour in the morning and the evening are all factors that will put one at a higher risk of getting into an accident. Human behavior such as alcohol consumption, over speeding and other activities are also important factors. The dangerous sections of the city were determined by the spatial analysis. Traffic prone areas, complicated road systems or low visibility characterize these hotspots. Temporal outcomes indicated that traffic control measures ought to be more centred on high seasons and time.

The results also revealed that pedestrians, cyclists, and motorcyclists are at a greater risk and the reasons are that more stringent safety laws and quality road infrastructure are necessary. The dataset was not without its flaws, though some aspects were not very good, including missing values and inconsistency of the fields. However, the outcomes provide a definite image of traffic accidents causes and how they are affected by various factors. Depending on the results, it is possible to make a number of recommendations: Enhance road design and road infrastructure in hotspots, e.g. placing better crossing points, better light and clearer signs. Building awareness among the companies that engage into speeding, distracted driving and night time driving hazards. Increase enforcement at the hours of the day when there is the maximum risk and in the area where there are frequent accidents. Install safer pathways to the walkers and bicycles such as sidewalks and exclusive bicycle lanes. Use data dashboards and real-time monitoring to allow authorities to identify and respond to the growing accident trends in good time. To sum up, it is possible to conclude that open data could be extremely helpful in the context of acquiring better insights into traffic accidents and assisting with the development of more effective safety measures. Through improved data, planning and educating the population, one can minimize the amount of accidents and make the roads a safer place.

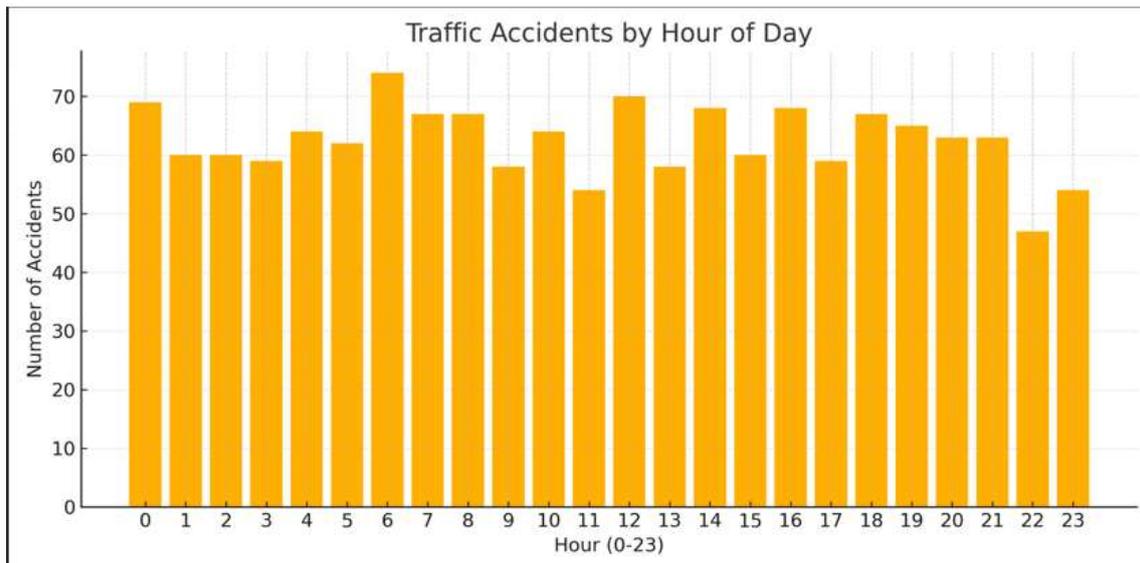


Figure 1 illustrates the frequency of the occurrence of the accidents on a traffic basis per hour of the day. The findings indicate clearly that there are two primary peaks one in the morning, and the other one in the evening.

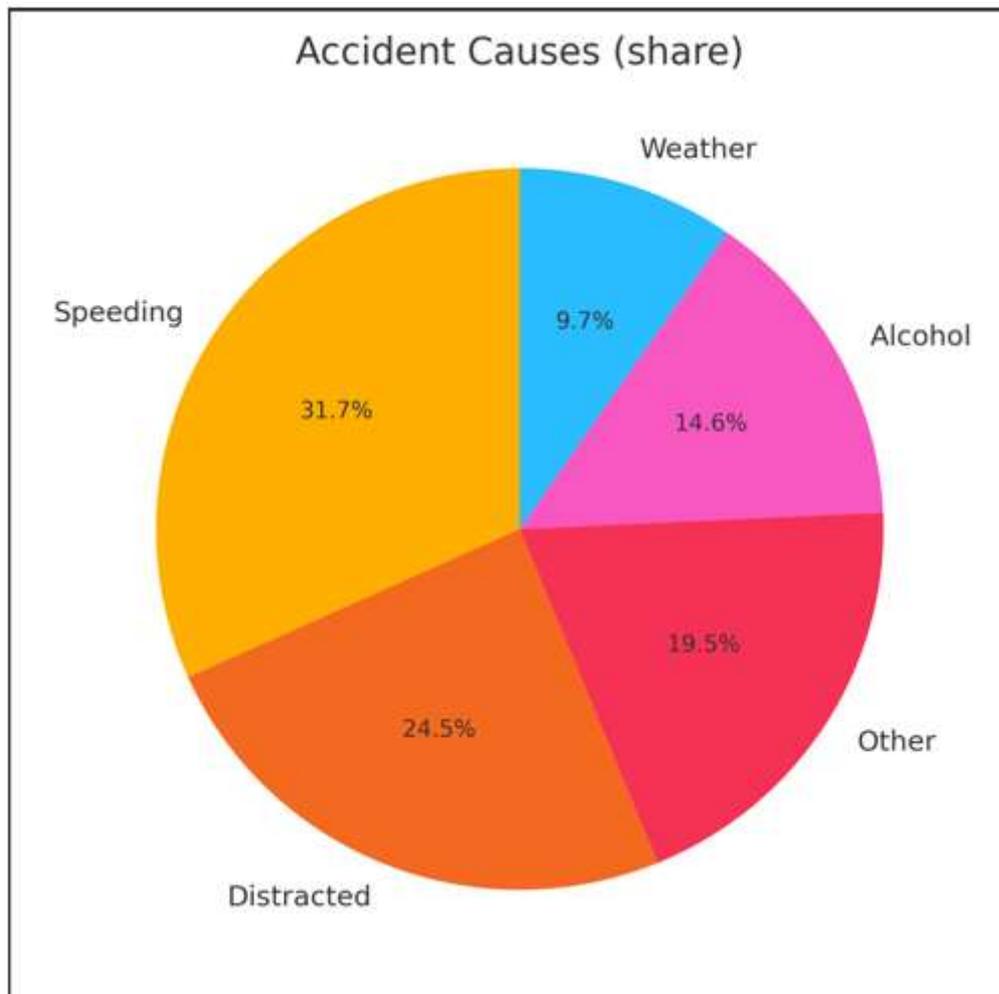
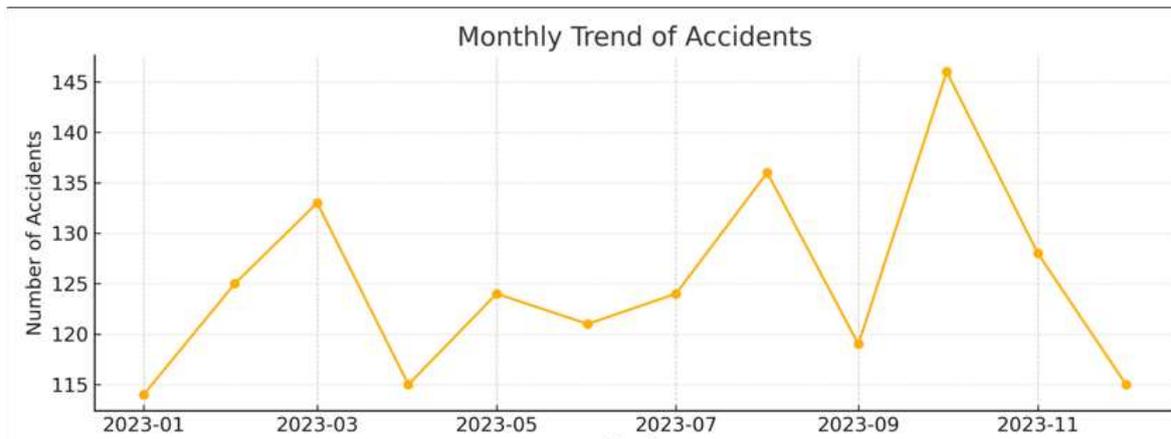


Figure 2 is a pie chart that illustrates the proportion of the various causes of accidents in the data set. Answers to speeding and distracted driving seem to be the most prevalent.



The changes in the number of accidents every month are displayed in Figure 3. The line chart shows that the number of accidents rises in the summer seasons and falls in winter.

### REFERENCES:

- Ainy, S., Soori, H., & Mahfozpour, S. (2011). Road traffic injuries and crashes: A review of international trends. *Journal of Transportation Safety*, 7(2), 45–52.
- Department for Transport. (2023). Reported Road Casualties in Great Britain: Annual Report. UK Government Open Data Portal. <https://data.gov.uk>
- European Commission. (2022). CARE: EU Road Accidents Database. European Road Safety Observatory. [https://ec.europa.eu/transport/road\\_safety](https://ec.europa.eu/transport/road_safety)
- National Highway Traffic Safety Administration. (2023). Fatality Analysis Reporting System (FARS). U.S. Department of Transportation. <https://www.nhtsa.gov>
- New York City Open Data. (2023). Motor Vehicle Collisions – Crashes. NYC Open Data Portal. <https://opendata.cityofnewyork.us>
- OECD International Transport Forum. (2020). Road Safety Annual Report. OECD Publishing.
- Toroyan, T. (2015). Global status report on road safety. World Health Organization (WHO).
- Yannis, G., Antoniou, C., & Papadimitriou, E. (2014). Impact of roadway, traffic, and environmental conditions on traffic accidents. *Accident Analysis & Prevention*, 71, 176–183.
- Ziakopoulos, A., & Yannis, G. (2020). A review of machine learning methods for crash prediction using open datasets. *Safety Science*, 135, 105–129.