Road Safety Analysis System

Group 25

Group members:

Kavya Gopal (6581-2209)

Kasiviswanathan Srikant Iyer (5222-2519)

Rema Veeranna Gowda (9462-8005)

Swaathi Reena Velavan (1230-8520)

Contents

Introduction

The following deliverable shows the conceptual design and user interface of our system. The conceptual design is shown through the help of an Entity Relationship Diagram. The various entities used in our proposed system are explained along with the important concepts used like primary key, multivalued attributes, composite attributes, etc. We have explained the relationships and the relationship cardinalities between various entities. We have also translated the entity relationship diagram into a relational schema. We have translated the entity relationship diagram into a relational schema taking into consideration the role of multivalued attributes, composite attributes, etc. This deliverable also includes a detailed user interface workflow which shows the various pages in the interface, how they are connected and the various components. Every section ends with a design rationale and a detailed explanation.

Overview of the system:

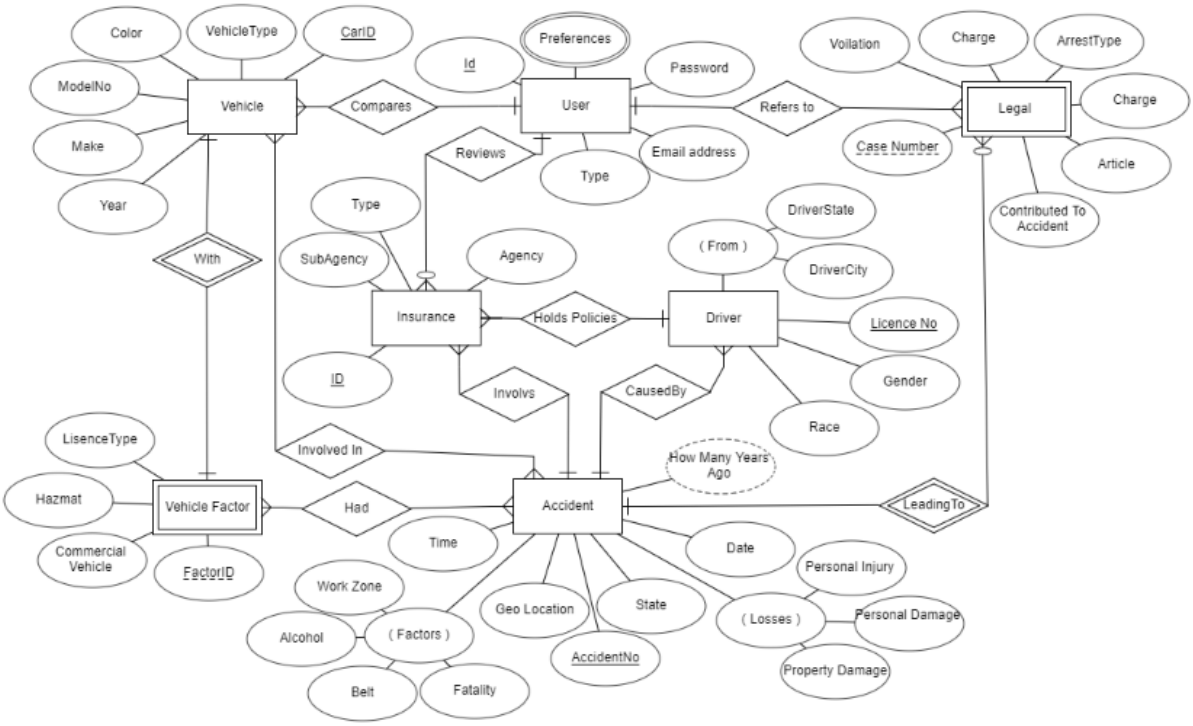
The motivation behind our design is to provide a platform for a variety of user roles (legal, insurance, car customers, health care workers) to analyse a comprehensive view of vehicle accidents. The real-world dataset that we have used for our project is based on the traffic violations events in Maryland County from 2012 to 2018. The dataset has 1.04 million tuple record. Our complex queries help users analyse the trends of accidents and make informed decisions. For example, car customers interested in a particular model and make, may investigate the trends of accident history over the years for the model. Another example is the usage of location trends (where accidents have predominantly occurred over the years) by health care workers for the purpose of resource planning. (where to ensure more ambulances/first aid kits/traffic men)

Analysis of the data source:

- The main table in the data source – “accident.csv” file has 1.04 million tuples of data with 15 columns.
- The Vehicle table from “vehicle.csv” file has details of all the vehicles that have been in an accident in these years. It has 5 columns.
- The user table has all the different types of users which access the data. This table has only as many tuples as the number of users who sign up to use the system. It has 5 columns and changing number of tuples.
- The legal.csv file has 7 columns with 0.9 million tuples.
- The driver.csv file has 6 columns with 1.03 tuples.
- Insurance.csv file has 4 columns with 300 tuples.
- Involved.csv file has relations between many of the above files and has 1.04 million tuples.
- This dataset was collected from https://www.data.gov/.

  The following is the link to the dataset –
  https://www.kaggle.com/rounak041993/traffic-violations-in-maryland-county

ER Diagram



Entities:

The project consists of the following entities:

1. **Vehicle** - This entity describes the vehicle that has partaken or been involved in an accident. It includes details like the vehicle type, colour, model number, make and car identification number. The car identification number forms the primary key of this entity.

2. **Accident** - Accident is an entity that describes when and where the accident has occurred. It is uniquely identified by the accident number attribute. The accident entity has other attributes like state and how many years ago this accident took place. This entity also has a composite attribute called losses.

   2.1. **Losses** -The composite attribute Losses is composed of atomic attributes like personal damage, personal injury and property damage. The attribute losses describe the losses incurred due to the accident. This entity has an attribute called “How many years ago”. This attribute is a derived attribute and is derived from the date of the accident.

   2.2. **Factors** - Factors is a composite attribute that describes the factors that played a role in the vehicle accident. This includes details like seat belt usage, alcohol consumption, fatality and whether it takes place in a work zone. These details play a role in the occurrence of any accident.

3. **Insurance** - This entity describes the insurance that is associated with supporting/compensating a particular accident. It has attributes that include Agency and Subagency and the primary key, ID.

4. **Driver -** This entity entails the driver involved in the cause of the accident. This includes attributes like Race, Gender and From. This entity is uniquely identified by the licence number attribute.

    4.1. The from attribute is a composite attribute that describes where the driver is from. It includes the composite attributes Driver city and Driver state.

5. **Legal -** This entity describes the legal issues surrounding an accident. It includes attributes like violation, charge, article, contributed to accident, arrest type, article, and case number. The case number is the primary key which uniquely identifies the entity.

6. **User -** This entity entails the user details that will use our system. It includes attributes like Id, password, User preferences, email address and the user type. The Id is used to uniquely identify the user; thus, it is the primary key.

    6.1. The attribute user preferences is a multivalued attribute.

7. **Vehicle Factor** – This entity is used to identify details about the type of vehicle. This entity has attributes such as licence type, Hazmat, commercial vehicle and factorID as the primary key.

Relationships:

- **Compares**: The user compares the vehicles they want to buy. Any one user can compare multiple vehicles at the same time. The cardinality of this relation will be 1: m.

- **With**: The vehicles have factors which may be a contributing factor to the accidents. Every vehicle will have a unique set of factors. As the factors cannot alone identify the vehicle, the vehicle factor is a weak entity. The cardinality of this relation will be 1:1.

- **Involved** in: The vehicles are involved in accidents. As many vehicles may be involved in an accident and one vehicle can be in multiple accidents. The cardinality of this relation will be m: n.

- **Leading** to: Any accident may lead to Law enforcement being involved. Any accident may be small enough to not involve the law enforcement and may be large enough to involve multiple cases. The cardinality of this relation will be 1:m where m can be 0.

- **Refers to**: The user may refer to the past records of cases present in Legal. The user may look up multiple issues at a time. The cardinality of this relation will be 1: m.

- **Reviews**: The user after their research may review the Insurance companies and rate them as per their preferences. A user can rate multiple insurance agencies. The cardinality of this relation will be 1:m where m can be 0.

- **Holds Policies**: Drivers are expected to hold Insurance policies in case of accidents. The driver can have multiple insurance policies but needs at least 1. The cardinality of this relation will be 1:m where m cannot be 0.

- **Caused By**: Any accident may be caused by one or many drivers. The cardinality of this relation will be 1:m where m cannot be 0.

- **Involves:** When an accident happens the insurance, companies are involved who the driver of the accident has policies with. Any accident will connect with every insurance company the driver has insurance with. The cardinality of this relation will be 1:m where m cannot be 0.

Relational Schema -

- **Vehicle**(CarID, Vehicle type, Color, ModelNo, Make, Year)
- **Vehicles_Factors**(FactorID, Belt, Fatality, Alcohol, Work Zone)
- **Accident**(Time , Date , GeoLocation, State , How many years ago , Personal Injury, Personal damage)
- **Insurance**(ID, Type, Agency, Subagency)
- **Driver**(LicenseNo, Race, Gender, Driver city, Driver state)
- **User**(Id, password, Email address, type, Preferences)
    - **User_Preferences** (Id, preferences)
- **Legal** (CaseNumber, Violations, Charges, Arrest type, Charges, Articles, Contribution to accident)

Tables:

1. **Vehicle**{CarID, Vehicle type, Color, ModelNo, Make, Year}

2. **Vehicles_Factors**{FactorID, Belt, Fatality, Alcohol, Work Zone, Car_ID, FOREIGN_KEY (CarID) REFERENCES (Vehicle)}

3. **Accident** { Time , Date , GeoLocation, State , How many years ago , Personal Injury, Personal damage, FOREIGN_KEY (CaseNumber) REFERENCES (Legal) }

4. **Insurance** {ID, Type, Agency, Subagency}

5. **Driver** {LicenseNo, Race, Gender, Driver city, Driver state}

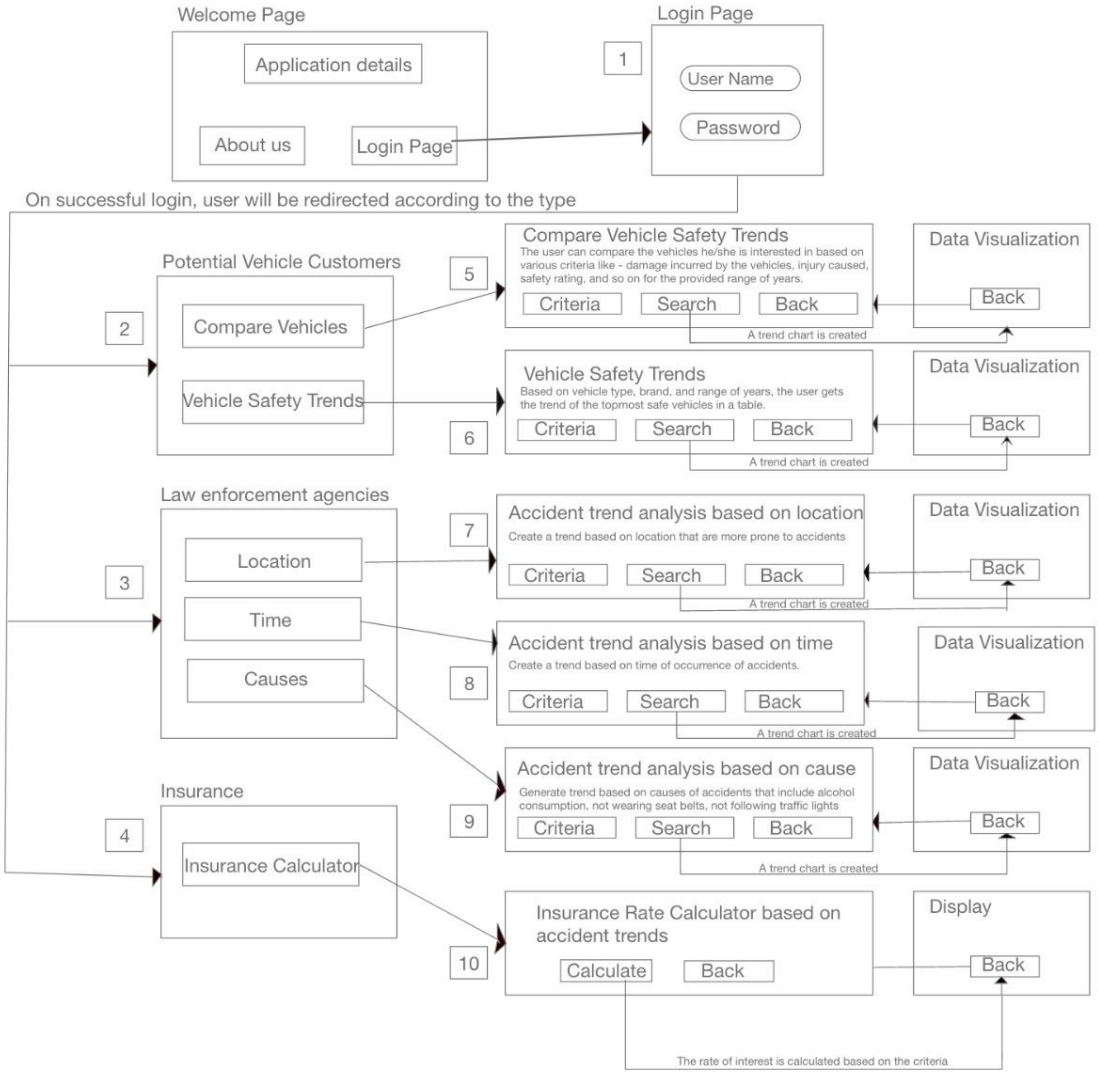
6. **User**{Id, password, Email address, type, Preferences, Insurance_ID, FOREIGN_KEY (ID as Insurance_ID) REFERENCES (Insurance)}

7. **Legal** {CaseNumber, Violations, Charges, Arrest type, Charges, Articles, Contribution to accident, AccidentNo, FOREIGN_KEY (AccidentNo) REFERENCES (Accident)}

8. **Involved** {InsuranceId, LisenceNo, CaseNumber, AccidentNo,CarID , FOREIGN_KEY (CarID) REFERENCES (Vehicle), FOREIGN_KEY (AccidentNo) REFERENCES (Accident), FOREIGN_KEY (CaseNumber) REFERENCES (Legal), FOREIGN_KEY (LisenceNo) REFERENCES (Driver), FOREIGN_KEY (InsuranceId) REFERENCES (Insurance)}

User interface Design and Application specific Network

Web Pages:



Application Specific Network of Web Pages

The user enters the username and password to login to webpage on opening the application. Based on the user type, the person is redirected to the user specific page. Criteria is given for the user to choose what trend the person wants to analyse like safety trend ,interest rate and so on. After the display it comes back to the normal page where it was called upon. The whole structure is designed in the form of a tree with each user specific trends. Each webpage will have logo of the site.

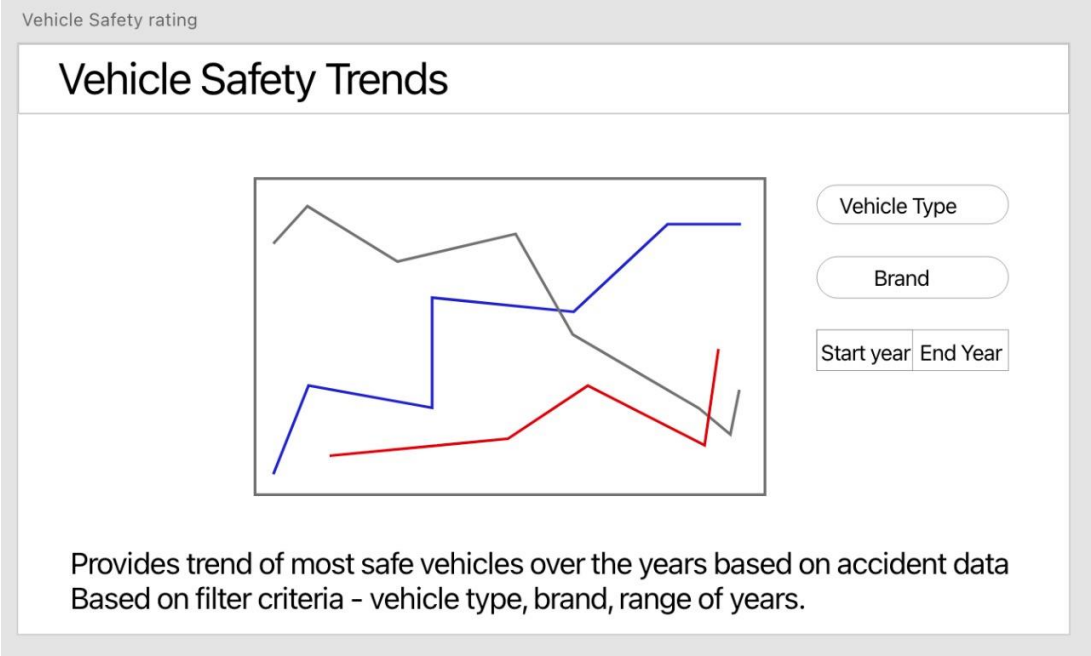
Welcome Page:

The user lands on the welcome page which has a brief overview of the applications of the road safely analysis system. The functionalities provided by this application are structured based on the type of users - Law enforcement agencies, health sectors, insurance companies, and potential vehicle buyers. The user can choose any one of these options to get detailed analysis and trends. For instance, clicking on “Vehicle Customer” would allow the user to access trends on the safest vehicles based on the accident data and also would help them compare among different vehicles on criteria like damage

caused during accidents, safety rating, injuries incurred, and so on. These insights would help the user to make an informed decision before buying a vehicle.

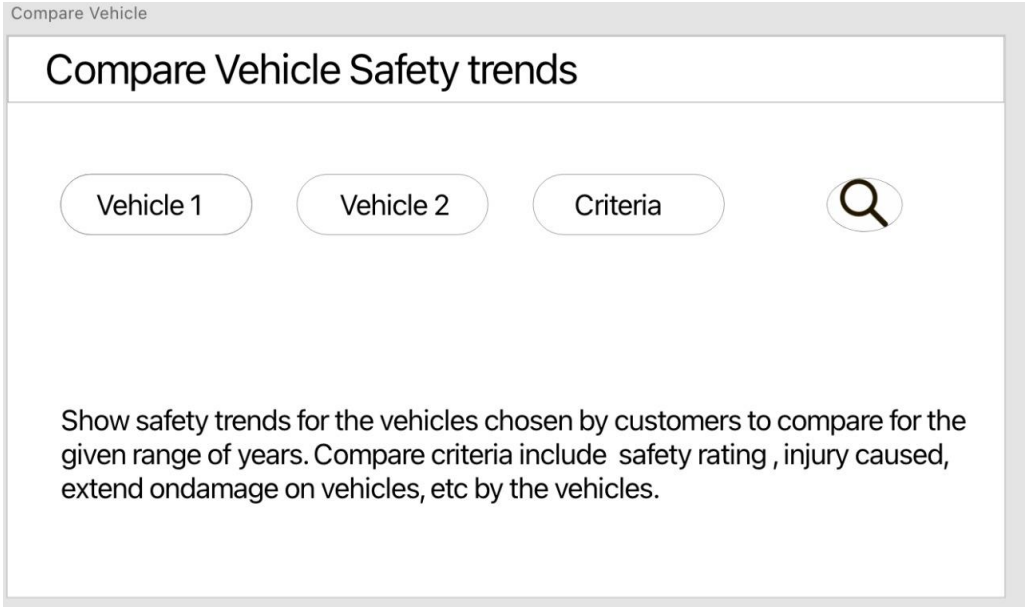
Potential Customers Page (2):

Potential customers who are looking to buy a vehicle are the primary target of this webpage where they can do the following:



- *Access Vehicle safely trends in order to make an informed decision before buying a vehicle:* The customers input the vehicle type, brand, and range of years to get the trend of the topmost safe vehicles over the specified range. For example, if a customer searches for the car by Toyota for the years 2000 to 2021, a graph is plotted which shows the accidents the vehicles have been a part of. The safest vehicle can be inferred from the graph by viewing the one that has the least number of accidents it has been involved in. (The line graph would have multiple lines for each car and hence provides the comparison in a visual format to the user). Also, the trend specifies if the safety features of the car have become better over the years and would give customers the confidence to buy it.

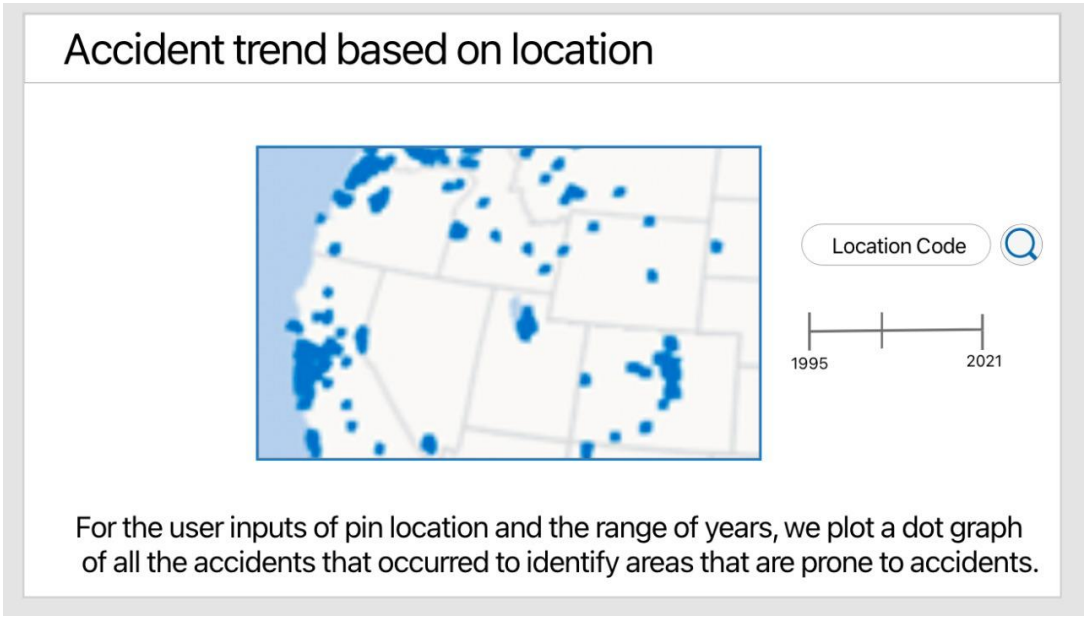
    - *Query:* Internally, we will query the DB to get tuples of vehicles that belong to the mentioned brand. For these tuples, we find the count of accidents the vehicle has been involved in from the accident table by performing join operation ("involved_in" table is used for this). We project only the tuples for the given year range. This is plotted as a multiline graph on the webpage to show the topmost safe vehicles.



- *Compare vehicles for safety standards by the mentioned criteria (6):* Often, the customers are confused between which vehicle to buy among the ones they like. The customer can use the application to compare the vehicles he/she is interested in based on various criteria like - damage incurred by the vehicles, injury caused, safety rating, and so on for the provided range of years.

Law Enforcement Page (3):

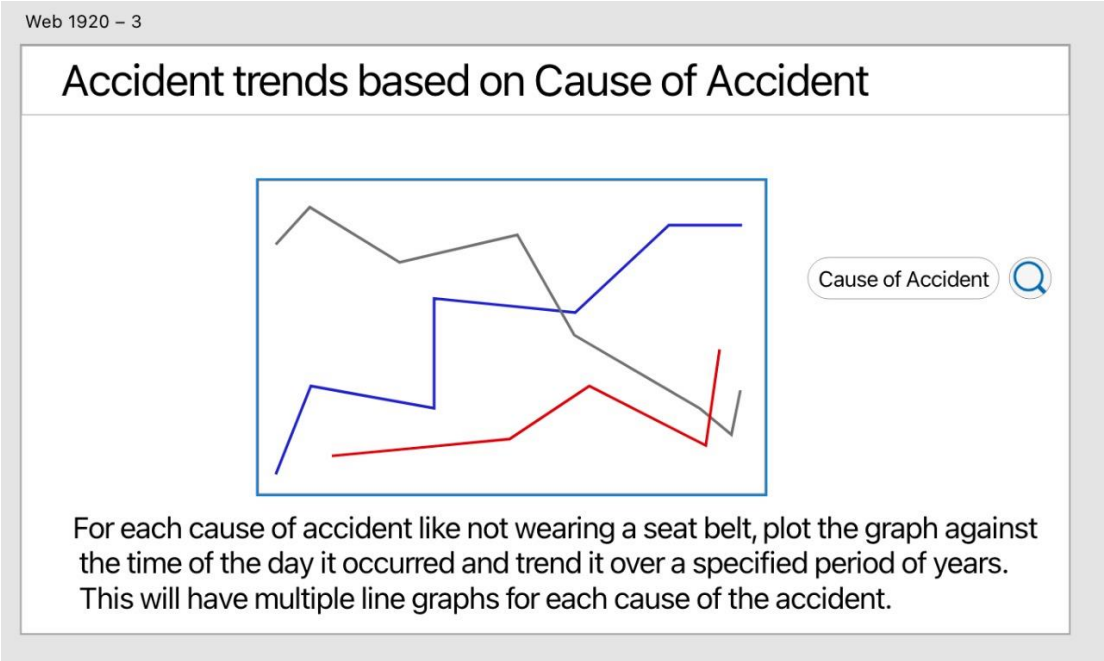
Primarily this webpage is designed for use by traffic engineers, policymakers and law enforcement agencies where they can do the following trend analysis to decide on corrective measures to reduce road accidents.



- *Accident Trend Analysis based on Location(7):* Often there are certain areas that are more prone to accidents and identifying these areas will help the law enforcement agencies to come up with reasonable solutions that would reduce the rate of accidents. The trends generated by the graph are dynamic with the year and location. After corrective measures are implemented

by the agencies, they can refer to the trends to see the impact of the measures taken as all trends over the range of years can be analysed.
Note: These trends can be used by the health sector to plan resources like placing ambulances in accident-prone areas. Also, it would help them target health promotions and injury prevention interventions on those locations.

- *Query:* The user inputs the pin location and gives the range of years for which we need the accident data to identify the accident-prone areas. Initially to plot a dot graph we filter the data based on year and location. We also provide a radius to query all the accidents that happened in and around that zip code of the area. The geological graph will have denser circles around areas where the count of accidents are high calculated from the accident table based on the attribute accident number and year is derived by using the system date and the date of the accident.



- *Accident Trend Analysis based on Cause of Accident(9):* Often the cause of accidents is something that could be controlled by having stricter law enforcement. For instance, some of the causes include alcohol consumption, not wearing seat belts, not following traffic lights, and so on. The user can input the causes of accidents to get the trend of times these accidents that happened over a period of time. Hovering over the graph gives a list of unique charges that was charged against these offenders. Firstly, the trend analysis would help the police and lawmakers identify the time during which most of these accidents happen and can plan monitoring activities accordingly. Secondly, getting the list of charges would help them analyze if there should be any change in the law that would be required which would reduce the accidents.

Insurance Page (4)

Insurance companies are the another type of user for the application as they can use the data analysed to set their insurance rates and make a better plan.

- *Insurance Rate Calculator(10):* The insurance. Companies will determine the interest to be charged from the user depending on the vehicle he wants to get insured. Every attribute of the vehicle will be taken into consideration to calculate the rate of interest by being sorted using

the features that will most probably cause accidents. Weightage of each attributes damage causing abilities will be added together find how dangerous the car is. For the car with highest probability of getting into an accident, highest possible interest will be charged while the safe cars will; be charged only the base interest rate.

- *Query:* Sort the accident table ascendingly using composite attribute Losses and join the table with Vehicle table to get the vehicles causing the greatest number of accidents using foreign key and then join with insurance company and create column 'interest rate' which is calculated based on the formula "Base Interest+ (additional interest*accident probability)"

Final Trend Queries:

1. *Access Vehicle safely trends in order to make an informed decision before buying a vehicle:* For the given user input of vehicle and vehicle brand, we count the number of accidents the type of vehicle has been in and plot it over a range of years. This is plotted as a multi-line graph on the web page to show the topmost safe vehicles with their safety trends.

2. *Compare vehicles for safety standards by the mentioned criteria:* For the given user inputs of vehicles, we compare the damage incurred by the vehicles during accidents, injury caused, safety rating, and so on for them and generate trends for the given range of years. We have line graphs denoting trends for each criterion.

3. *Accident Trend Analysis based on Location:* For the user inputs of pin location and the range of years, we plot a dot graph of all the accidents that occurred to identify areas that are prone to accidents

4. *Accident Trend Analysis based on Time:* For the user inputs of pin location and the range of years to generate a trend analysis.

5. *Accident Trend Analysis based on Cause of Accident:* For each cause of accident like not wearing a seat belt, plot the graph against the time of the day it occurred and trend it over a specified period of years. This will have multiple line graphs for each cause of the accident.

Work Division

Below is the work division on each phase of the product.

Database: Kasiviswanathan Srikant Iyer, Kavya Gopal
PHP (backend): Rema Veeranna Gowda, Kasiviswanathan Srikant Iyer
Website frontend: Swaathi Reena Velavan, Rema Veeranna Gowda
Queries: Kavya Gopal, Swaathi Reena Velevan

For every consecutive phase the responsibilities will be passed on in a cyclic manner so that everyone gets to work on every part of the project and will be registered on the documentation submitted further.