

# **PROJECT MILESTONE 1**

## **PROJECT DETAIL:**

- **Name of the project:** Analysis of US Accidents 2021 dataset
- **Team:** Level Up
- **Members:**
  - a. Prathamesh Kakade (50460522)
  - b. Shravani Soma (50477925)
  - c. Sriinitha Chinnapatlola (50478024)

## **PROBLEM STATEMENT:**

### **DESCRIPTION OF THE DATA:**

This is a countrywide car accident dataset, which covers 49 states of the USA. The accident data are collected from February 2016 to Dec 2021, using multiple APIs that provide streaming traffic incident (or event) data. These APIs broadcast traffic data captured by a variety of entities, such as the US and state departments of transportation, law enforcement agencies, traffic cameras, and traffic sensors within the road-networks. Currently, before any preprocessing is applied, there are about 2.8 million accident records in this dataset.

### **DESCRIPTION OF PROBLEM:**

The objective is to design a relational database system that will use existing “accidents in US” data to provide insightful analytics and help existing Department of Transportation, as well as the future employees of the department to keep track and make important decisions regarding expansion, construction, and maintenance of roads in the United States.

An exponential amount of data is created per day on the internet. Storing huge amounts of data and retrieving knowledge from it is a challenging task. This data will surely help in gaining meaningful insights of what conditions and reasons could be the result of an accident and how it can be prevented for the safety of every individual. It provides an opportunity to the Department of Transportation to improve their services and for the Managers and employees working within to choose the best possible way out to help reduce any road accidents.

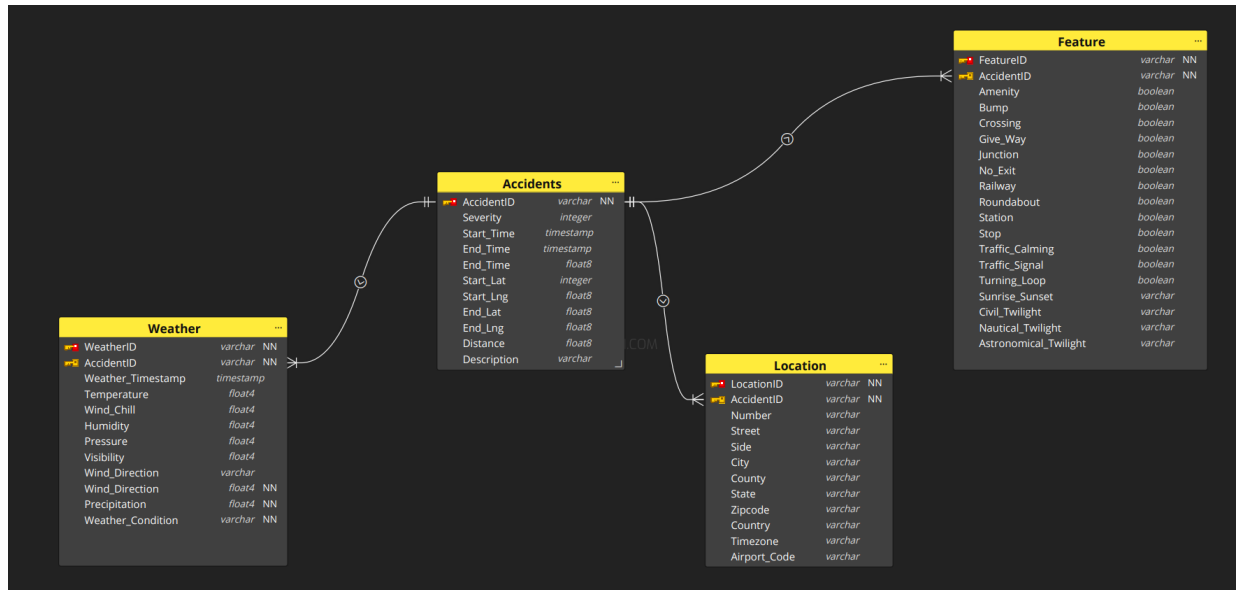
### **TARGET USER AND ADMIN:**

Users of the system will be the managers who are willing to keep track of the mishaps that happen on the road and make important decisions based on the visual plots such as expanding the roads, making sure they are well maintained or building bridges and tunnels that will help in reducing the number of road accidents.

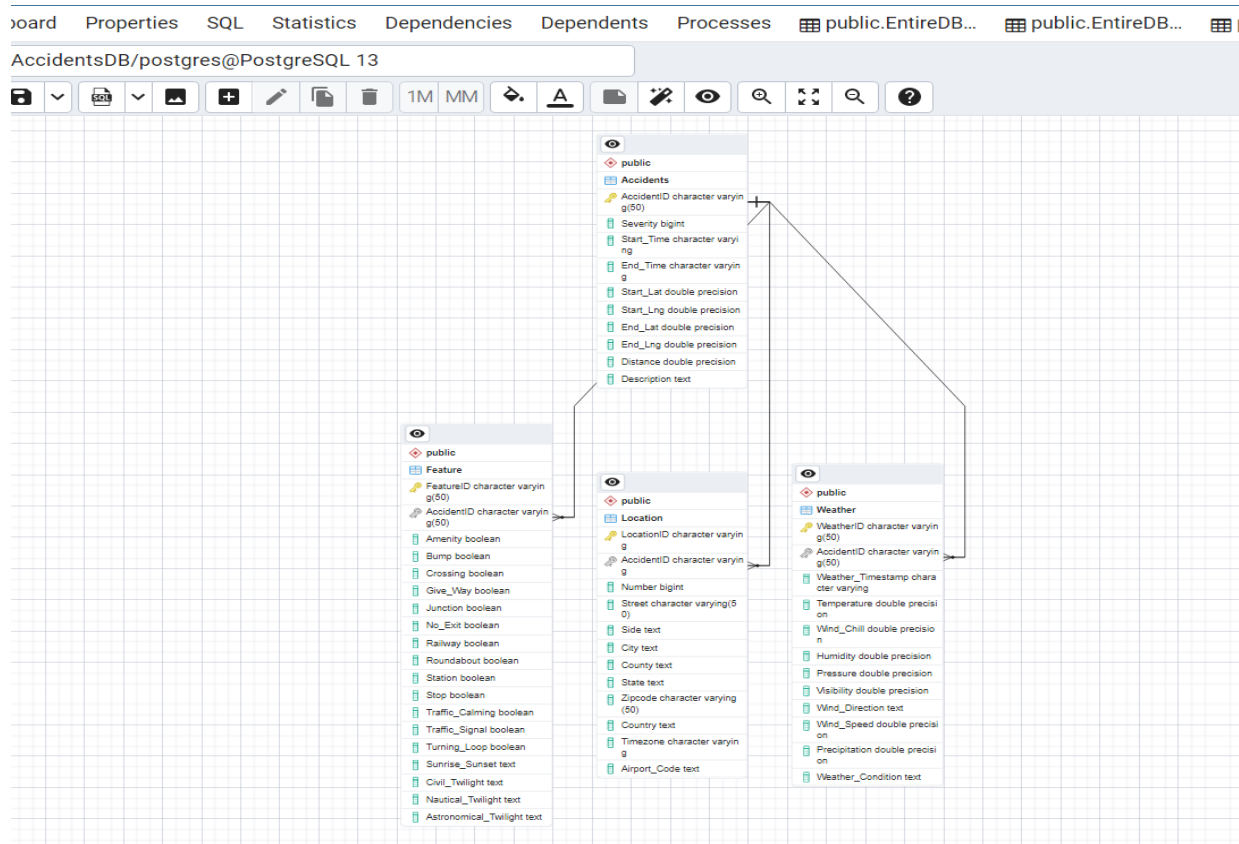
The Administrator of the system will be an employee of IT sector in the Department of Transportation. They will be responsible of maintaining this system and making sure all new data entries are added, modified, or required old entries are deleted from the database system. The admin will have the responsibility to check that the database is running (as we will be building UX for the database system as well).

## E/R Diagram:

## Created Manually:



## Created Using PostgreSQL:



## **RELATIONSHIP BETWEEN TABLES:**

- a. The Weather table is linked to Accidents table using AccidentID.
- b. The Location table is linked to Accidents table using AccidentID.
- c. The Feature table is linked to Accidents table using AccidentID.

## **DATA VISUALIZATIONS:**

To justify the problem statement and to understand various relations in the data, we have plotted few graphs that reflect the information in the dataset pictorially.

1. Accidents with respect to the Latitudes and Longitudes in the dataset:

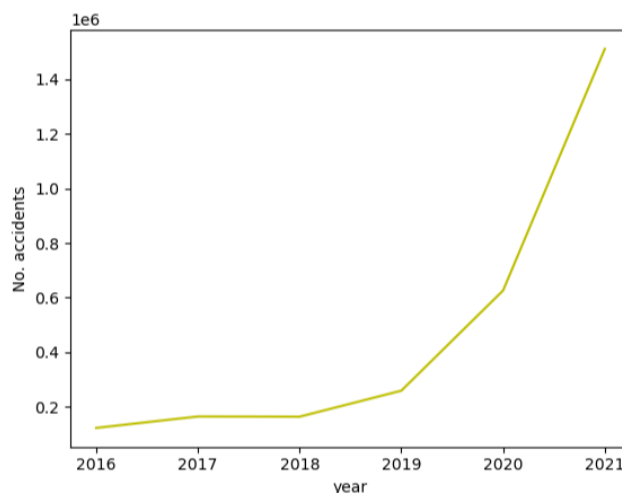
This scatterplot is plotted using the Latitudes and longitudes given in the dataset. It can be inferred that there are numerous accidents happening throughout the country (USA).

```
<AxesSubplot:xlabel='Start_Lng', ylabel='Start_Lat'>
```

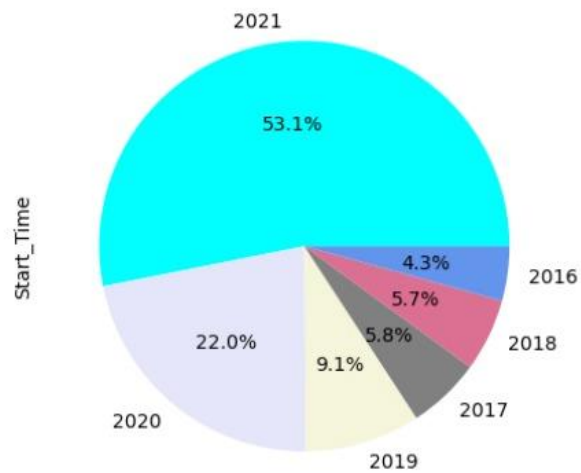


### **2. Accident Trends over the years (2016 – 2021):**

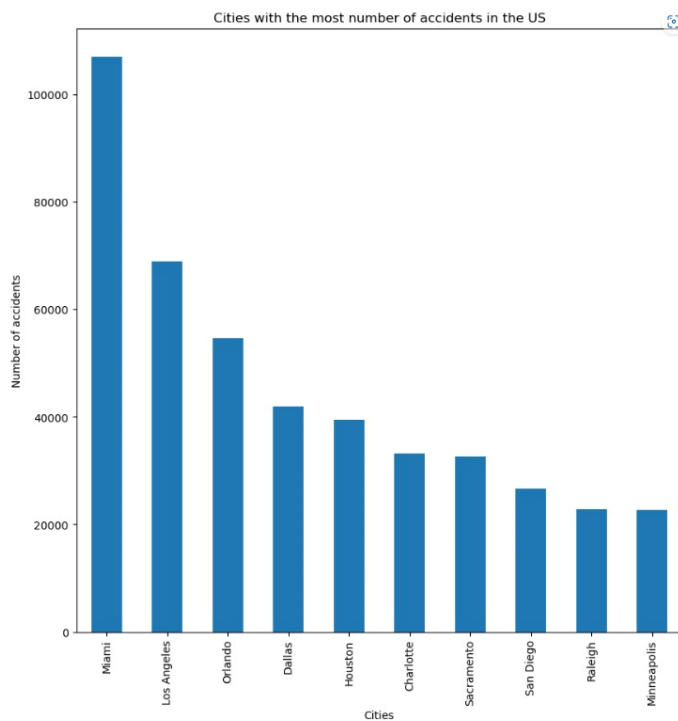
The dataset has accident-related data from the year 2016 until 2021. Here's a line graph representing the trend in the number of accidents over the years. It is observed that the number of accidents has been increasing over the years with a sudden increase particularly observed in between 2020-2021.



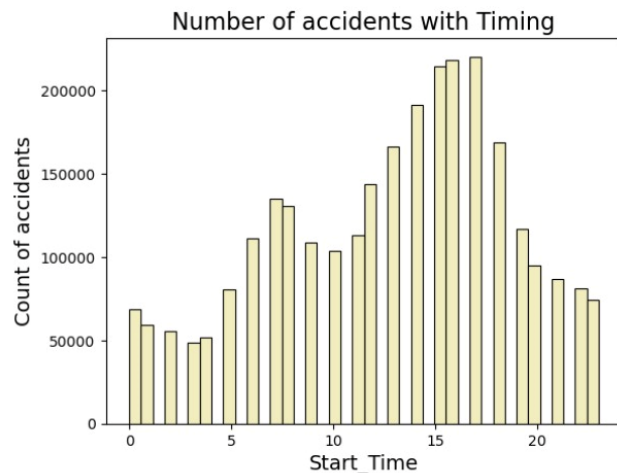
**3. Pie Chart:** Here's a pie chart describing the percentage of accidents in each year (in the dataset). It can be inferred that most of the accidents occurred in the year 2021 which agrees with the above line graph.



**4. Cities with the highest number of accidents:** There are many cities in the dataset, to know the cities with the highest number of accidents over the years we plotted a bar graph that shows the top 10 cities with highest accident rates. From the plot, it can be inferred that Miami has the highest number of accidents recorded over the years.



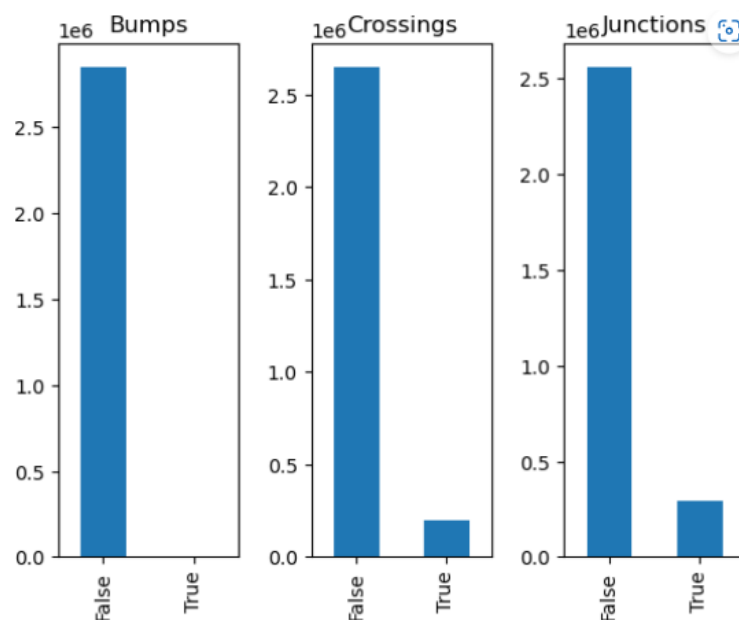
**5. Accident rate vs time of the day:** A histogram has been plotted to find out the time at which maximum accidents occurred. And from the below plot, we can infer that the greatest number of accidents are occurred around 3pm - 5pm which makes sense as most the people commute from workplace during that time.



## 6. Subplots to analyze the relationship between the number of accidents and road conditions like bumps, crossings, and junctions:

From the subplot below, the following inferences can be made:

- It can be inferred that all the three parameters played a very minimal role in the occurrence of accidents. The value “True” means that there was a bump/crossing/junction near to the place where an accident occurred. The value “False” means that there was no bump/crossing/junction near to the place where an accident occurred.
- It seems like there’s minimal to no effect of bumps on the number of accidents.



## **DATABASE IMPLEMENTATION:**

### **DATA SCHEMAS**

#	Attribute	Description
1	ID	This is a unique identifier of the accident record.
2	Severity	Shows the severity of the accident, a number between 1 and 4, where 1 indicates the least impact on traffic (i.e., short delay as a result of the accident) and 4 indicates a significant impact on traffic (i.e., long delay).
3	Start_Time	Shows start time of the accident in local time zone.
4	End_Time	Shows end time of the accident in local time zone. End time here refers to when the impact of accident on traffic flow was dismissed.
5	Start_Lat	Shows latitude in GPS coordinate of the start point.
6	Start_Lng	Shows longitude in GPS coordinate of the start point.
7	End_Lat	Shows latitude in GPS coordinate of the end point.
8	End_Lng	Shows longitude in GPS coordinate of the end point.
9	Distance(mi)	The length of the road extent affected by the accident.
10	Description	Shows natural language description of the accident.
11	Number	Shows the street number in address field.
12	Street	Shows the street name in address field.
13	Side	Shows the relative side of the street (Right/Left) in address field.

#	Attribute	Description
14	City	Shows the city in address field.
15	County	Shows the county in address field.
16	State	Shows the state in address field.
17	Zipcode	Shows the zipcode in address field.
18	Country	Shows the country in address field.
19	Timezone	Shows timezone based on the location of the accident (eastern, central, etc.).
20	Airport_Code	Denotes an airport-based weather station which is the closest one to location of the accident.
21	Weather_Timestamp	Shows the timestamp of weather observation record (in local time).
22	Temperature(F)	Shows the temperature (in Fahrenheit).
23	Wind_Chill(F)	Shows the wind chill (in Fahrenheit).
24	Humidity(%)	Shows the humidity (in percentage).
25	Pressure(in)	Shows the air pressure (in inches).
26	Visibility(mi)	Shows visibility (in miles).
27	Wind_Direction	Shows wind direction.
28	Wind_Speed(mph)	Shows wind speed (in miles per hour).
29	Precipitation(in)	Shows precipitation amount in inches, if there is any.

#	Attribute	Description
30	Weather_Condition	Shows the weather condition (rain, snow, thunderstorm, fog, etc.)
31	Amenity	A POI annotation which indicates presence of amenity in a nearby location.
32	Bump	A POI annotation which indicates presence of speed bump or hump in a nearby location.
33	Crossing	A POI annotation which indicates presence of crossing in a nearby location.
34	Give_Way	A POI annotation which indicates presence of give_way in a nearby location.
35	Junction	A POI annotation which indicates presence of junction in a nearby location.
36	No_Exit	A POI annotation which indicates presence of no_exit in a nearby location.
37	Railway	A POI annotation which indicates presence of railway in a nearby location.
38	Roundabout	A POI annotation which indicates presence of roundabout in a nearby location.
39	Station	A POI annotation which indicates presence of station in a nearby location.
40	Stop	A POI annotation which indicates presence of stop in a nearby location.
41	Traffic_Calming	A POI annotation which indicates presence of traffic_calming in a nearby location.
42	Traffic_Signal	A POI annotation which indicates presence of traffic_signal in a nearby location.



#	Attribute	Description
43	Turning_Loop	A POI annotation which indicates presence of turning_loop in a nearby location.
44	Sunrise_Sunset	Shows the period of day (i.e. day or night) based on sunrise/sunset.
45	Civil_Twilight	Shows the period of day (i.e. day or night) based on civil twilight.
46	Nautical_Twilight	Shows the period of day (i.e. day or night) based on nautical twilight.
47	Astronomical_Twilight	Shows the period of day (i.e. day or night) based on astronomical twilight.

### **RELATIONSHIP BETWEEN TABLES:**

- a. The Weather table is linked to Accidents table using AccidentID.
- b. The Location table is linked to Accidents table using AccidentID.
- c. The Feature table is linked to Accidents table using AccidentID.

### **ATTRIBUTES**

#### **Accident Relation:**

Name	Data type
AccidentID (PK)	varchar (50)
Severity	int
Start_Time	text
End_Time	text
Start_Lat	decimal (9,6)
Start_Lng	decimal (9,6)
End_Lat	decimal (9,6)
End_Lng	decimal (9,6)
Distance	decimal (10,2)
Description	text

**Weather Relation:**

Name	Data type
WeatherID (PK)	varchar (50)
AccidentID (FK)	varchar (50)
Weather_Timestamp	text
Temperature	decimal (9,6)
Wind_Chill	decimal (9,6)
Humidity	decimal (9,6)
Pressure	decimal (9,6)
Visibility	decimal (9,6)
Wind_Direction	Text
Precipitation	decimal (9,6)
Weather_Condition	Text

**Location Relation:**

Name	Data type
LocationID	varchar (50)
AccidentID	int
Number	varchar (50)
Street	text
Side	text
City	text
County	text
State	text
Zipcode	varchar
Country	text
Timezone	varchar

**Feature Relation:**

Name	Data type
FeatureID	varchar (50)
AccidentID	varchar (50)
Amenity	boolean
Bump	boolean
Crossing	boolean
Give_Way	boolean
Junction	boolean
No_Exit	boolean
Railway	boolean
Roundabout	boolean
Station	boolean
Stop	boolean
Traffic_Calming	boolean
Traffic_Signal	boolean
Traffic_Loop	boolean
Sunrise_Sunset	text
Civil_Twilight	text
Nautical_Twilight	text
Astronomical_Twilight	text

# INSERTION OF DATA IN POSTGRESQL:

## Entire Database:

Data OutputMessagesNotifications

	ID text	Severity bigint		Start_Time text	End_Time text	Start_Lat double precision	Start_Lng double precision	End_Lat double precision	End_Lng double precision	Distance double precision	
1	A-43		4	2016-02-09 18:20:58	2016-02-10 00:20:58	40.45112	-85.15048	40.35429	-85.14993	6.0	
2	A-44		4	2016-02-09 18:20:58	2016-02-10 00:20:58	40.35429	-85.14993	40.45112	-85.15048	6.0	
3	A-48		4	2016-02-10 06:18:49	2016-02-10 12:18:49	40.72813	-84.78965	40.74559	-84.78962	1.20	
4	A-51		2	2016-02-10 08:35:27	2016-02-10 14:35:27	41.83193	-80.10143000000002	41.84149	-80.11099	0.8240000000000000	
5	A-67		2	2016-02-10 12:54:39	2016-02-10 18:54:39	41.48339	-81.66297	41.47692	-81.66075	0.40	
6	A-90		2	2016-02-11 07:20:03	2016-02-11 13:20:03	38.33667	-81.65623000000002	38.33614	-81.65623000000002	0.03700000000000000	
7	A-91		2	2016-02-11 07:20:03	2016-02-11 13:20:03	38.33614	-81.65623000000002	38.33667	-81.65623000000002	0.03700000000000000	
8	A-113		2	2016-02-11 13:30:58	2016-02-11 19:30:58	40.58919	-80.09885	40.58919	-80.09885		
9	A-119		2	2016-02-11 16:56:28	2016-02-11 22:56:28	40.58919	-80.09885	40.58919	-80.09885		
10	A-149		3	2016-02-13 07:14:41	2016-02-13 13:14:41	40.484222	-80.13755400000002	40.503456	-80.139196	1.30	
11	A-192		4	2016-02-15 20:46:40	2016-02-16 02:46:40	38.824929	-85.47449499999998	38.82415	-85.63794	8.70	
12	A-204		2	2016-02-16 06:08:42	2016-02-16 12:08:42	41.06347	-81.50372	41.06472	-81.50414	0.08900000000000000	
13	A-264		2	2016-02-17 17:04:42	2016-02-17 23:04:42	41.47395	-81.69931	41.47865	-81.6931	0.40	
14	A-438		4	2016-02-24 12:27:56	2016-02-24 18:27:56	41.427584	-85.8495	41.471378	-85.839527	3.00	
15	A-439		4	2016-02-24 12:27:57	2016-02-24 18:27:57	41.471378	-85.839527	41.427584	-85.8495	3.00	
16	A-462		4	2016-02-25 06:14:20	2016-02-25 12:14:20	39.97527	-85.14018	39.98511	-85.14406	0.00	
17	A-463		4	2016-02-25 06:14:20	2016-02-25 12:14:20	39.98511	-85.14406	39.97527	-85.14018	0.00	
18	A-476		4	2016-02-25 13:13:48	2016-02-25 19:13:48	39.9672	-81.28699999999998	39.96392	-81.27197	0.80	
19	A-477		4	2016-02-25 13:13:48	2016-02-25 19:13:48	39.96392	-81.27197	39.9672	-81.28699999999998	0.80	
20	A-12532		2	2016-12-07 07:47:35	2016-12-07 13:47:35	38.814617	-104.75764	38.81371	-104.75765	0.00	

Total rows: 1000 of 943318Query complete 00:00:23.124Ln 1, Col 1

### Queries for Table creation:

```

1 cur.execute('''CREATE TABLE Accident (
2     AccidentID varchar(50) PRIMARY KEY,
3     Severity int,
4     Start_Time timestamp,
5     End_Time timestamp,
6     Start_Lat decimal(9,6),
7     Start_Lng decimal(9,6),
8     End_Lat decimal(9,6),
9     End_Lng decimal(9,6),
10    Distance decimal(10,2),
11    Description varchar(255));''')
12 print("Table created successfully")
13 conn.commit()
14
15 cur.execute('''CREATE TABLE Feature (
16     FeatureID varchar(50) PRIMARY KEY,
17     AccidentID varchar(50),
18     Amenity boolean,
19     Bump boolean,
20     Crossing boolean,
21     Give_way boolean,
22     Junction boolean,
23     No_Exit boolean,
24     Railway boolean,
25     Roundabout boolean,
26     Station boolean,
27     Stop boolean,
28     Traffic_Calming boolean,
29     Traffic_Signal boolean,
30     Turning_Loop boolean,
31     Sunrise_Sunset varchar(10),
32     Civil_Twilight varchar(10),
33     Nautical_Twilight varchar(10),
34     Astronomical_Twilight varchar(10),
35     FOREIGN KEY (AccidentID) REFERENCES Accidents(AccidentID)
36 );''')
37 print("Table created successfully")
38 conn.commit()
39
40 cur.execute('''CREATE TABLE Heather (
41     HeatherID varchar(50) PRIMARY KEY,
42     AccidentID varchar(50),
43     Weather_Timestamp datetime,
44     Temperature_F decimal(5,2),
45     Wind_Chill_F decimal(5,2),
46     Humidity decimal(5,2),
47     Pressure decimal(6,2),
48     Visibility_mi decimal(5,2),
49     Wind_Direction varchar(10),
50     Wind_Speed_mph decimal(5,2),
51     Precipitation_in decimal(5,2),
52     Weather_Condition varchar(50),
53     FOREIGN KEY (AccidentID) REFERENCES Accidents(AccidentID)
54 );''')
55 print("Table created successfully")
56 conn.commit()
57
58 cur.execute('''CREATE TABLE Feature2 (
59     FeatureID2 int PRIMARY KEY,
60     AccidentID int,
61     Amenity boolean,
62     Bump boolean,
63     Crossing boolean,
64     Give_way boolean,
65     Junction boolean,
66     No_Exit boolean,
67     Railway boolean,
68     Roundabout boolean,
69     Station boolean,
70     Stop boolean,
71     Traffic_Calming boolean,
72     FOREIGN KEY (AccidentID) REFERENCES Accidents(AccidentID)
73 );''')
74
75 conn.commit()
76
77
78 conn.close()

```

### Accident Table:

[illegible]

Location Table:

Data OutputMessagesNotifications

	LocationID [PK] character varying	AccidentID character varying	Number bigint	Street character varying (50)	Side text	City text	County text	State text	Zipcode character varying (50)
1	L-1	A-1	9001	W State Road 26	R	Dunkirk	Jay	IN	47336
2	L-10	A-10	5	Forest Grove Rd	L	Coraopolis	Allegheny	PA	15108-3485
3	L-100	A-100	4836	Highway 212 E	L	Chaska	Carver	MN	55318-9249
4	L-1000	A-1000	804	20th Ave S	R	Minneapolis	Hennepin	MN	55454
5	L-10000	A-10000	2014	Caton Ave	R	Brooklyn	Kings	NY	11226-2804
6	L-100000	A-100000	16390	Biscayne Blvd	L	North Miami Beach	Miami-Dade	FL	33160
7	L-100001	A-100001	10072	Cottonwood Rd	R	Bozeman	Gallatin	MT	59718-8969
8	L-100002	A-100002	781	Sumneytown Pike	R	Lansdale	Montgomery	PA	19446-5301
9	L-100003	A-100003	344	Monroe Ave	L	Rochester	Monroe	NY	14607-3662
10	L-100004	A-100004	901	Cedar St	L	Flint	Genesee	MI	48503-3657
11	L-100005	A-100005	1098	Washington St	R	Waukegan	Lake	IL	60085-5429
12	L-100006	A-100006	300	N Davis St	L	Jacksonville	Duval	FL	32202-4816
13	L-100007	A-100007	2633	Ridge Ave	L	Dayton	Montgomery	OH	45414-5431
14	L-100008	A-100008	9990	NC Highway 62 N	R	Milton	Caswell	NC	27305-9346
15	L-100009	A-100009	60792	Highway 101	L	Coos Bay	Coos	OR	97420
16	L-10001	A-10001	75906	US Highway 101	L	Reedsport	Douglas	OR	97467
17	L-100010	A-100010	1803	W Broadway St	R	Oviedo	Seminole	FL	32765
18	L-100011	A-100011	1831	Old Hardin Rd	L	Billings	Yellowstone	MT	59101-6562
19	L-100012	A-100012	3180	College Dr	R	Baton Rouge	East Baton Rouge	LA	70808-3119
20	L-100013	A-100013	3701	Cahuenga Blvd	R	Studio City	Los Angeles	CA	91604-3504

Total rows: 1000 of 943318

Query complete 00:00:07.212

Successfully run. Total query runtime: 7 secs 212 msec. 943318 rows affected.

Feature Table:

Data OutputMessagesNotifications

	FeatureID [PK] character varying (50)	AccidentID character varying (50)	Amenity boolean	Bump boolean	Crossing boolean	Give_Way boolean	Junction boolean	No_Exit boolean	Railway boolean	Roundabout boolean	Station boolean	Stop boolean
1	F-1	A-1	false	true	true	true	false	false	false	true	false	true
2	F-10	A-10	false	false	false	true	true	true	false	false	true	false
3	F-100	A-100	false	false	true	false	false	true	true	false	true	true
4	F-1000	A-1000	false	true	true	false	false	true	true	true	true	true
5	F-10000	A-10000	true	true	false	false	true	true	false	false	true	false
6	F-100000	A-100000	false	false	false	true	true	true	false	true	false	true
7	F-100001	A-100001	true	true	false	true	true	false	true	true	true	true
8	F-100002	A-100002	false	false	true	true	true	false	false	true	false	false
9	F-100003	A-100003	true	true	true	true	true	false	false	false	true	false
10	F-100004	A-100004	false	true	true	false	true	true	false	true	false	false
11	F-100005	A-100005	true	true	false	true	false	false	true	false	false	false
12	F-100006	A-100006	true	false	false	false	true	true	false	false	true	true
13	F-100007	A-100007	false	true	true	false	true	true	true	false	false	true
14	F-100008	A-100008	false	true	false	true	true	false	false	true	true	false
15	F-100009	A-100009	false	true	true	true	true	true	true	false	true	false
16	F-10001	A-10001	false	false	false	true	true	true	true	true	true	true
17	F-100010	A-100010	false	false	false	false	false	true	true	true	false	false
18	F-100011	A-100011	false	true	true	false	false	false	false	true	false	true
19	F-100012	A-100012	false	false	true	true	true	false	true	true	true	true
20	F-100013	A-100013	false	false	true	false	true	true	true	true	true	false

Total rows: 1000 of 943318

Query complete 00:00:08.588

Successfully run. Total query runtime: 8 secs 588 msec. 943318 rows affected.

## Weather Table:

Data Output Messages Notifications											
	WeatherID [PK] character varying (50)	AccidentID character varying (50)	Weather_Timestamp character varying	Temperature double precision	Wind_Chill double precision	Humidity double precision	Pressure double precision	Visibility double precision	Wind_Direc text		
1	W-1	A-1	2016-02-09 18:20:00	19.9	7.3	81	29.85	2	WNW		
2	W-10	A-10	2016-02-13 06:51:00	6.1	-12.2	63	30.32	10	WNW		
3	W-100	A-100	2017-01-25 06:53:00	30	20.5	88	29.7	1.2	North		
4	W-1000	A-1000	2017-01-30 09:07:00	25	13.1	85	29.67	10	South		
5	W-10000	A-10000	2021-11-06 11:51:00	49	49	44	30.33	10	CALM		
6	W-100000	A-100000	2021-12-12 14:53:00	82	82	58	30.12	10	ENE		
7	W-100001	A-100001	2021-11-21 14:56:00	46	46	30	25.75	10	WNW		
8	W-100002	A-100002	2021-12-05 17:55:00	47	43	58	29.95	10	SE		
9	W-100003	A-100003	2021-10-18 17:54:00	48	46	80	29.39	10	W		
10	W-100004	A-100004	2021-11-16 18:53:00	39	33	60	29.23	10	SE		
11	W-100005	A-100005	2021-07-21 11:55:00	69	69	75	29.39	10	NNE		
12	W-100006	A-100006	2021-05-24 14:53:00	87	87	43	30.14	10	SE		
13	W-100007	A-100007	2021-09-17 17:56:00	86	86	43	29.03	10	SW		
14	W-100008	A-100008	2021-12-31 06:53:00	54	54	97	29.35	0.5	CALM		
15	W-100009	A-100009	2021-09-23 07:56:00	53	53	100	30.22	10	N		
16	W-10001	A-10001	2021-11-30 21:56:00	51	51	96	30.34	5	CALM		
17	W-100010	A-100010	2021-10-04 17:53:00	83	83	63	29.89	10	E		
18	W-100011	A-100011	2021-04-02 17:53:00	71	71	12	26.23	10	SW		
19	W-100012	A-100012	2021-07-22 15:20:00	87	87	67	30.04	10	S		
20	W-100013	A-100013	2021-10-07 01:51:00	63	63	82	29.14	10	S		

Total rows: 1000 of 943318 Query complete 00:00:06.034

✓ Successfully run. Total query runtime: 6 secs 34 msec. 943318 rows affected. ✕










## PRIMARY AND FOREIGN KEYS

- Accidents Relation
  - Primary key - AccidentID
  - Foreign key - No Foreign keys
- Weather Relation
  - Primary key - WeatherID
  - Foreignkey - id referenced from Accidents relation on Accidents.AccidentID
  - Foreignkey - id referenced from Location relation on Location.LocationID
- Location Relation
  - Primary key - LocationID
  - Foreignkey - id referenced from Accidents relation on Accidents.AccidentID
- Feature Relation
  - Primary key - FeatureID
  - Foreignkey - id referenced from Accidents relation on Accidents.AccidentID

## SAMPLE QUERIES:









### 1. Display Severity of all Accidents in Erie County:

```
Select "Accidents"."Severity"  
from "Accidents"  
join "Location" on "Accidents"."AccidentID" = "Location"."AccidentID"  
where "Location"."County"='Erie';
```

Data Output Messages Notifications		
       		
	Severity bigint 	
1	4	
2	4	
3	2	
4	2	
5	2	
6	2	
7	2	
8	4	
Total rows: 17 of 17 Query complete 00:00:00.700		

### 2. Display Description of all Accidents with AccidentID between A-100 and A-120:










```
Select "Description"  
from "Accidents"  
where "AccidentID"  
between 'A-100' and 'A-120';
```

Data Output Messages Notifications		
       		
	Description text	
1	Closed between US-421 and IN-3 - Road closed due to accident.	
2	Ramp to Exit 125B - Accident.	
3	At CR-43 - Accident.	
4	At Cologne - Accident.	
5	At Broadway St - Accident.	
6	At Galpin Blvd - Accident.	
7	At 31st St - Accident. Right lane closed.	
8	At 31st St - Accident.	
9	At CR-21/Trout Lake Rd - Accident.	
10	Closed at CO-12/Highway of Legends - Road closed due to accident.	
11	Closed at US-87/US-85/I-25-BL/Main St/7th St - Road closed due to accident.	
12	At 208th Ln - Accident.	
13	Closed at Frank Dr - Road closed due to accident.	
14	At Calhoun St - Accident.	
15	At Iron Horse Trl - Accident.	
16	At NE-98/854th Rd - Accident.	



3. **Display Count of all Accidents that had a Junction nearby:**












```
Select COUNT("Junction")
from "Feature"
where "Junction"='True';
```

Data Output		Messages	Notifications
       			
	<b>count</b> bigint 		
1	472348		

Total rows: 1 of 1    Query complete 00:00:00.633

4. **Display AccidentID, Wind Speed and Humidity for AccidentID=A-1000:**

```
Select "Accidents"."AccidentID", "Weather"."Wind_Speed", "Weather"."Humidity"
from "Weather"
join "Accidents"
on "Weather"."AccidentID"="Accidents"."AccidentID"
where "Accidents"."AccidentID"='A-1000';
```

Data Output				Messages	Notifications
       					
	<b>AccidentID</b> character varying (50) 	<b>Wind_Speed</b> double precision 	<b>Humidity</b> double precision 		
1	A-1000	13.8	85		
Total rows: 1 of 1    Query complete 00:00:00.224					

## **REFERENCE**

Dataset: <https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents>

Papers:

- Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. "A Countrywide Traffic Accident Dataset.", 2019.
- Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. "Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights." , 2019.

Other References:

- <https://learn.g2.com/data-preprocessing#:~:text=The%20four%20stages%20of%20data%20preprocessing%201%201.,3.%20Data%20reduction%20...%204%204.%20Data%20transformation>
- <https://www.postgresqltutorial.com/postgresql-python/connect/>
- <https://www.geeksforgeeks.org/sql-using-python/>

### **Contribution:**

Team Member	Milestone 1 part	Contribution (%)
Prathamesh kakade	Data collection, preprocessing, queries, report	33.33%
Shravani soma	Tables creation, Loading the data into pgAdmin, report	33.33%
Sriinitha Chinnapatlola	ER diagram, Visualizations, report	33.33%