Traffic Safety Project 2022:

Data Analysis and Recommendations for Reducing Road Accidents

National Highway Traffic Safety Administration (NHTSA)






Munzir Arsyuddin (Group Lead) #1, Dwiki Priambodo #2, Friendki #3, Sofia Octafiani Dewi #4, Erni Meilani #5, Nazrudin Ahwari #6

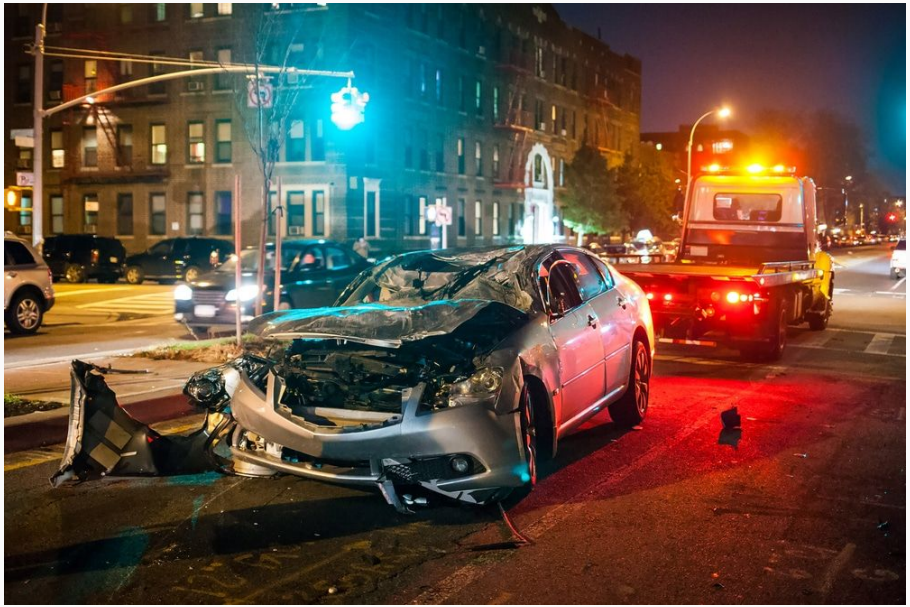
Background

The US National Highway Traffic Safety Administration (NHTSA) is the agency responsible for road traffic safety in the United States, receiving annual reports on car accidents resulting in thousands of fatalities and injuries.

Objective

To find effective ways to reduce road traffic accident rates in the upcoming year (2022) using traffic accident data from the year 2021.

Problem Statement



Identifying the factors that led to accidents during the year 2021.

Identify factors that cause accidents

1. Identification of Factors Leading to Accidents
2. Conditions that increase the risk of accidents.
3. Top 10 states with the highest number of accidents.
4. Average number of accidents per day based on the time of day accidents occurred.
5. Percentage of accidents caused by drunk drivers.
6. Percentage of accidents in rural and urban areas.
7. Number of accidents by day.
8. Intersection with driver condition per hour.

Data Cleaning and Data Cleaning Limitations

Data Cleaning Limitations

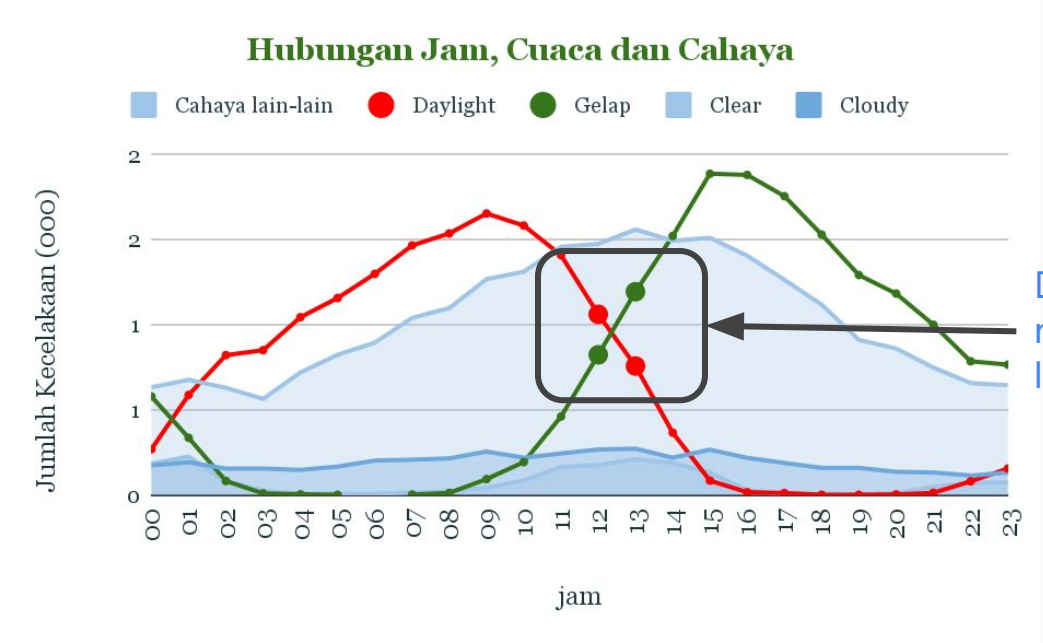
Comprehensive Data Examination

Query for Comprehensive Data Examination

```
SELECT
  to_char(waktulokal,'HH24') jam,-- mengambil waktu hanya jam dengan format 24H
    CASE
      -- menggabungan cahaya yang bukan terang dan gelap menjadi satu "Cahaya lain-lain"
      WHEN light_condition_name IN ('Dusk','Dawn','Reported as Unknown','Other','Not Reported')
        THEN 'Cahaya lain-lain'
      -- menggabungkan kondisi langit gelap menjadi gelap
      WHEN light_condition_name IN ('Dark - Lighted','Dark - Not Lighted','Dark - Unknown Lighting')
        THEN 'Gelap'
      ELSE light_condition_name
    END cahaya,
    CASE
      --menggabungjan yang bukan clear, rain,cloudy, dan not reported menjadi cuaca lain-laon
      WHEN atmospheric_conditions_1_name IN ('Fog, Smog, Smoke','Snow','Reported as Unknown',
      'Severe Crosswinds','Freezing Rain or Drizzle','Blowing Snow','Sleet or Hail','Other',
      'Blowing Sand, Soil, Dirt')
        THEN 'Cuaca Lain-lain'
      ELSE atmospheric_conditions_1_name
  END cuaca,
  COUNT(consecutive_number) jumlah_kecelakaan
FROM new_crash
GROUP BY jam,cahaya,cuaca
```

Comprehensive Data Examination

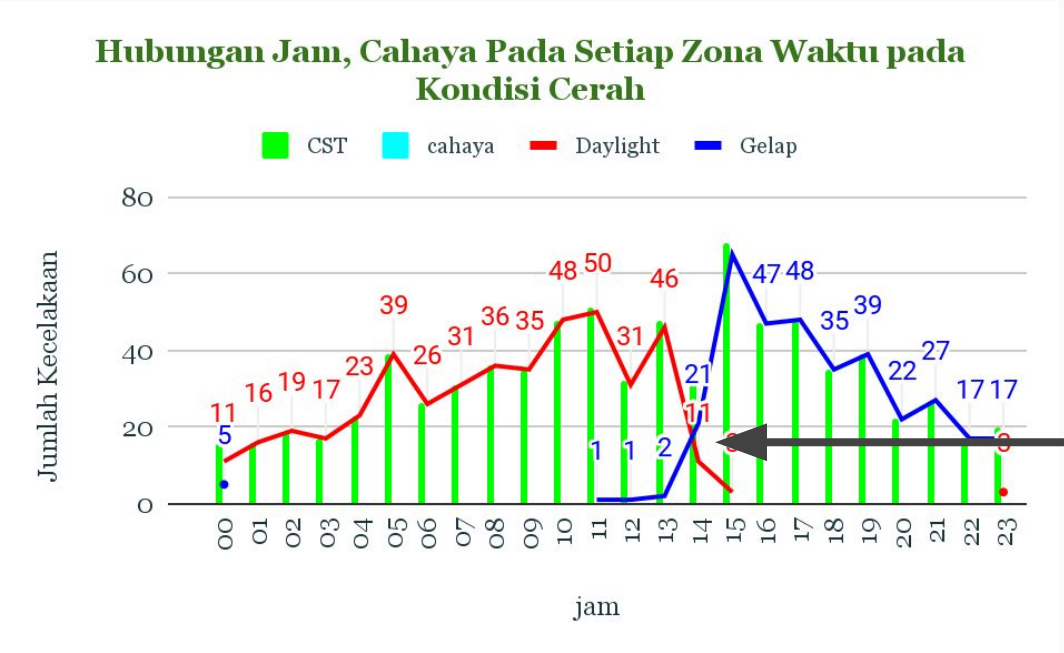
Visualization of Query Results



Data with almost the same number of accidents between light and dark

Data Examination by Month and Time Zone

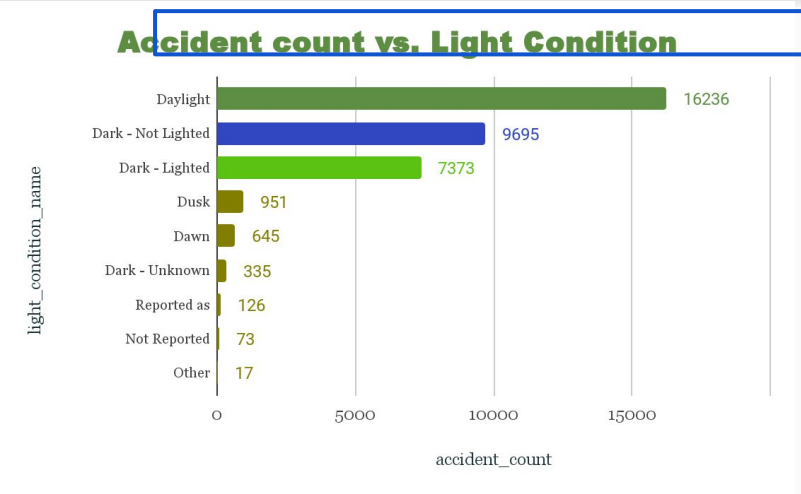
Visualization of Query Results

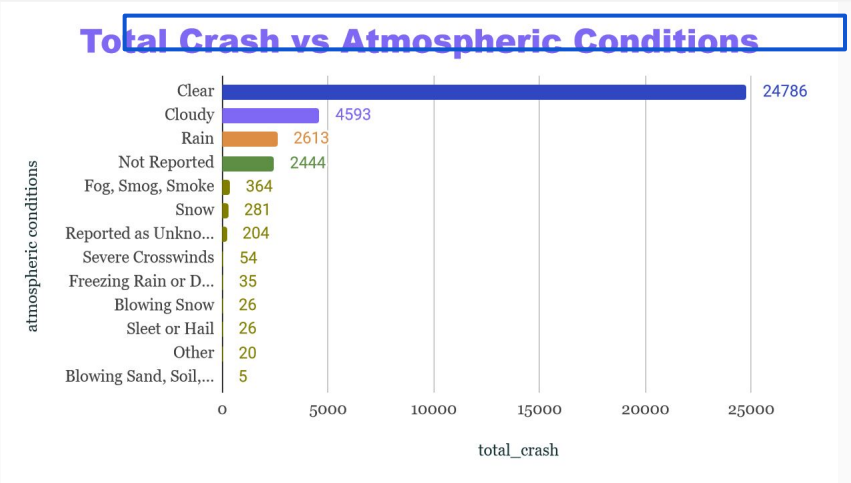


There is data with a 2:1 ratio between dark and light

Conditions That Increase Accidents

Conditions That Increase Accidents

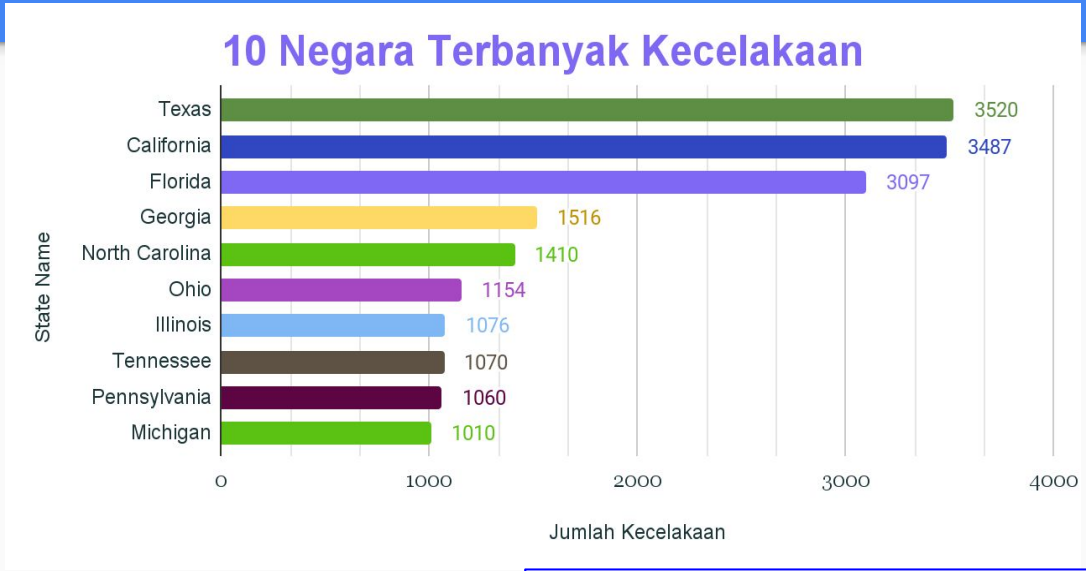




Sunny (Clear) Atmosphere Conditions & Daytime Lighting Conditions (Daylight) These are the atmospheric conditions where most accidents occur.
This is an important finding that suggests other factors may be more influential in causing accidents.

Top Ten States with Most Accidents

Visualisasi & Insight



There were 18,400 (51.9%) of the total 35,451 accidents that occurred in the 10 countries with the most accidents.

The states of Texas, California and Florida accounted for 28.5% of the total accidents with accident cases exceeding 3000.

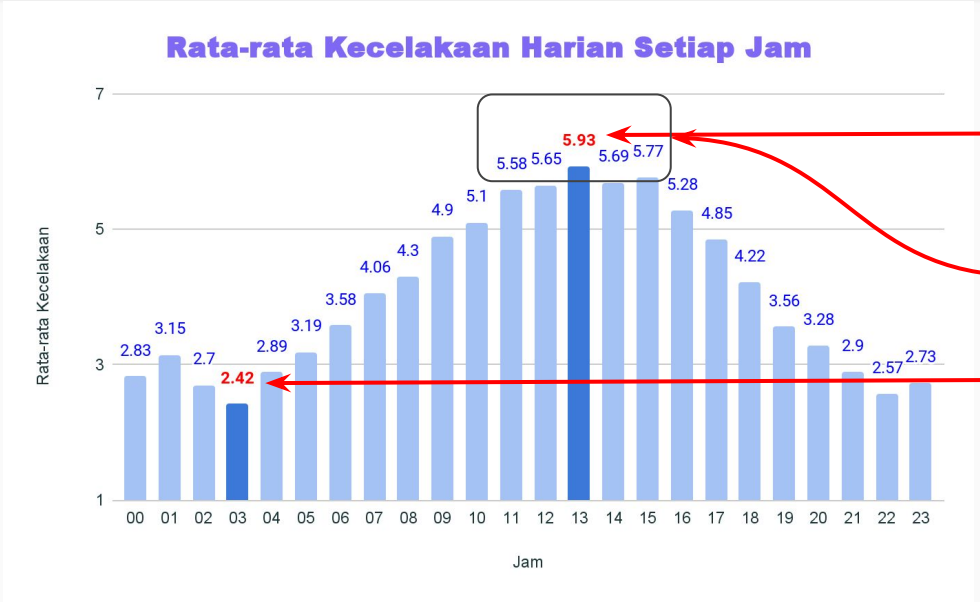
Recommendation

Focus on accident control on the 10 countries with the most because they account for 51.9% of the number of accidents

Average Accident Events per Day based on Time of Accident Occurrence

Average accidents per day by hour

Visualization of query results



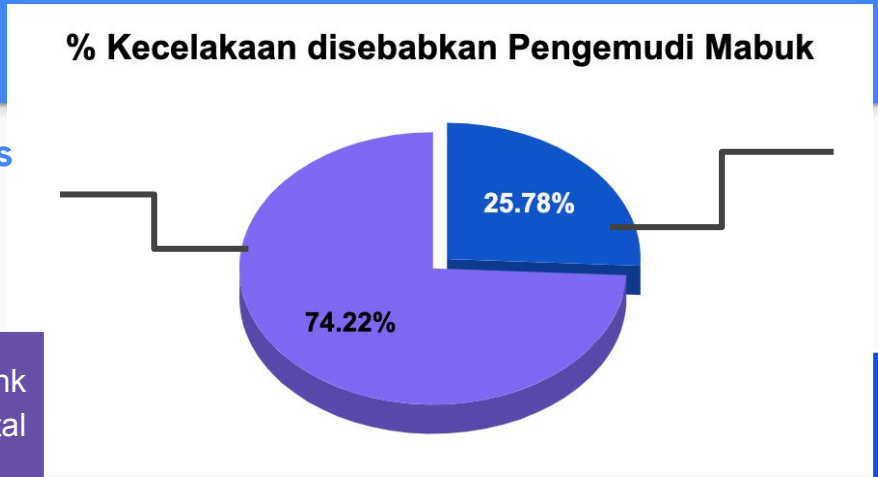
The highest average accident occurred at 13 o'clock

Accidents with an average above 5.5 occurred at 11-15 hours

The lowest average accident occurs at 03 hours

Percentage of Accidents caused by Drunk Driver

Accident Analysis Based on Drunk Drivers



Percentage of accidents caused by non-drunk drivers

Countif (number of drunk driver ,"=0") / count (total kecelakaan) * 100

Total accidents: 26310

Percentage of accidents caused by drunk drivers

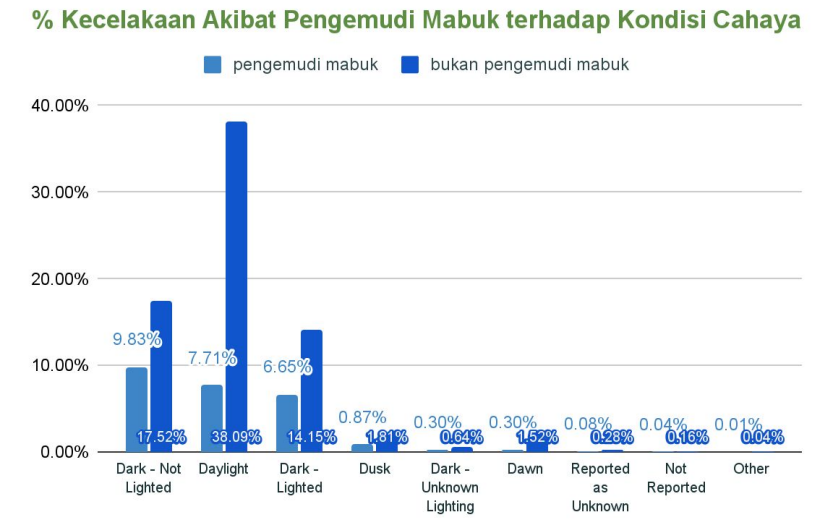
Countif (number of drunk driver ,">0") / count (total kecelakaan) * 100

Total accidents: 9141

The total number of accidents that occurred in 2021 was 35,451 accident cases

- number_of_drunk_drivers column to calculate the number of accidents caused by drunks or not

% of Accidents Due to Drunkenness Influenced by Light Conditions

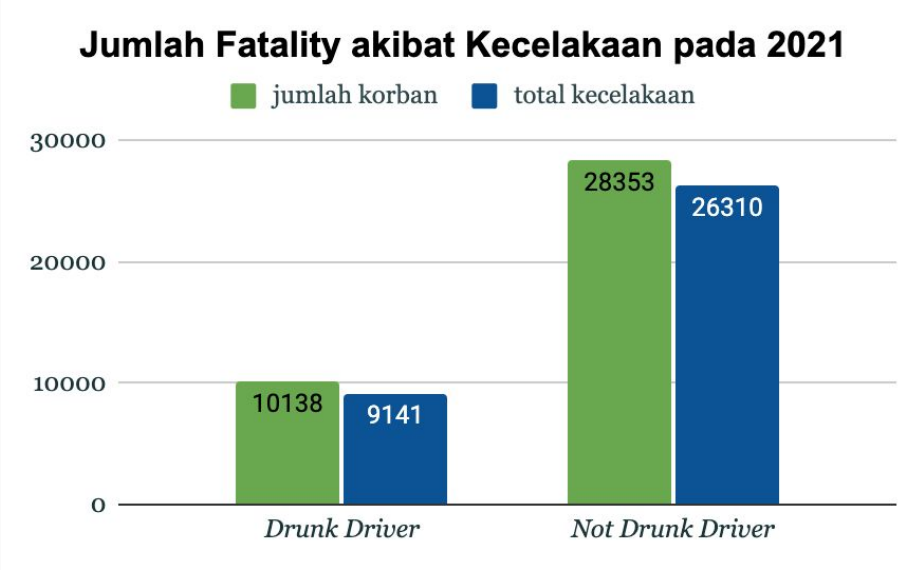


detail

Light Condition	Drunk Driver	Not Drunk Driver
Dark - Not Lighted	9.83%	17.52%
Daylight	7.71%	38.09%
Dark - Lighted	6.65%	14.15%
Dusk	0.87%	1.81%
Dark - Unknown Lighting	0.30%	0.64%
Dawn	0.30%	1.52%
Reported as Unknown	0.08%	0.28%
Not Reported	0.04%	0.16%
Other	0.01%	0.04%
Grand Total	**25.78%**	**74.22%**

- **Dark - not lighted** is a condition where accidents due to drunkenness occur most often
- **Daylight** is a condition where accidents caused by non-drinkers most often occur

Number of Fatalities in 2021 Based on Drunk or Not Drunk Drivers



Drunk driver

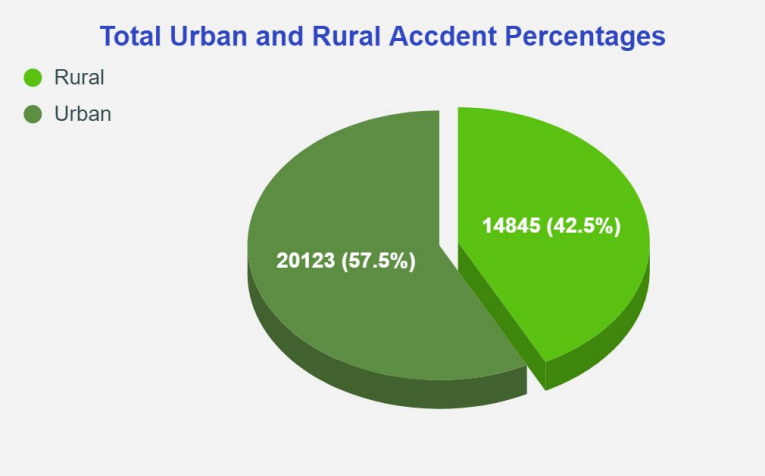
There were 10138 fatalities from accidents caused by drunk drivers.

Not Drunk driver

There were 28353 fatalities from accidents caused by non-drunk drivers.

Accidents in Rural and Urban Areas

Accidents in Rural and Urban Areas



Accident in Urban

Total accidents in urban areas 20123 with a percentage of 57.5%

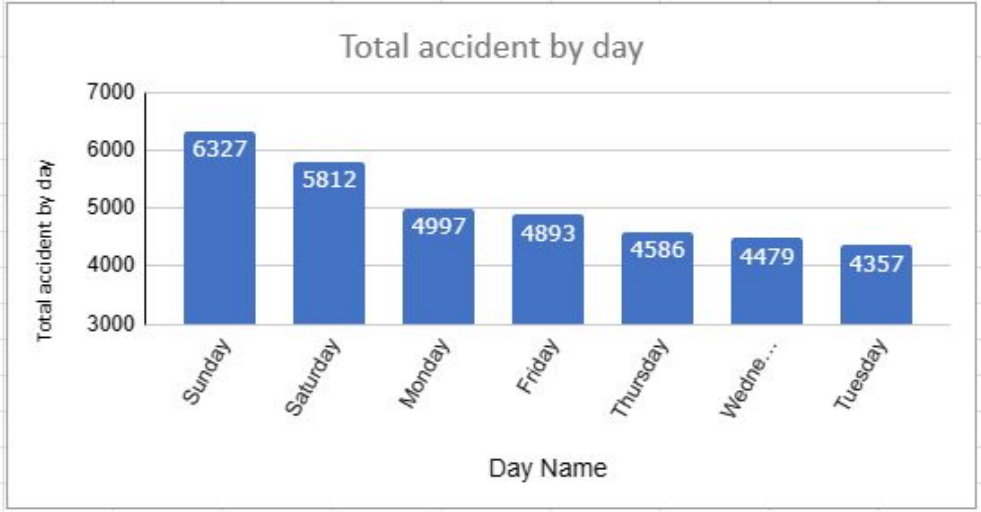
Accident in Rural

And the total number of accidents in rural areas was 1485 with a percentage of 42.5%.

Number of Accidents by Day

Query Results & Data Visualization of Number of Accidents by Day

Day Name	Total accident by day
Sunday	6327
Saturday	5812
Monday	4997
Friday	4893
Thursday	4586
Wednesday	4479
Tuesday	4357

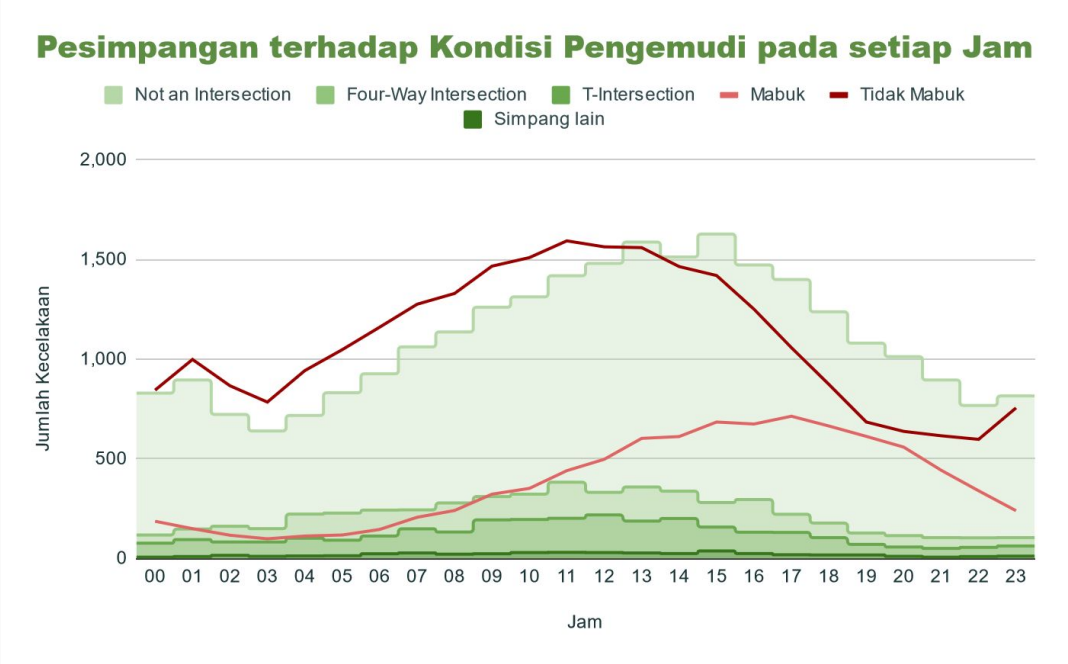


The highest number of accidents occurred on weekends, namely Sunday with 6327 cases and Saturday with 5812 cases. Meanwhile, on weekdays the number of accidents was <5000 cases, the lowest number was on Tuesday with 4357 cases

Intersections to Driver Conditions per Hour

Intersections against driver conditions per hour

Visualization of query results



The condition of drunk drivers is not the majority of conditions when accidents occur

Accidents occur on average on straight roads (without intersections) every hour

Between 18-20 hours the causes of accidents are balanced between drunk and not drunk conditions

CONCLUSION

1. **Light Conditions Review**

   Reevaluate light condition definitions for more accurate analysis.

2. **Driver Education**

   Promote safe driving habits during clear weather and daylight.

3. **Infrastructure Lighting**

   Improve lighting on poorly lit road sections for better visibility.

4. **Avoid DUI (Driving Under the Influence)**

   Raise awareness about the dangers of driving while impaired.