DATA ANALYSIS AND RECOMMENDATIONS FOR REDUCING ROAD ACCIDENTS
NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION (NHTSA)

MELLIZA NASTASIA IZAZI

ANALYSIS OF TRAFFIC ACCIDENTS IN THE UNITED STATES IN 2021

BACKGROUND

THE US NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION (NHTSA), A GOVERNMENT DEPARTMENT FOCUSED ON REDUCING TRAFFIC ACCIDENTS ON HIGHWAYS. IN RESPONSE TO THE HIGH NUMBER OF TRAFFIC ACCIDENTS IN THE UNITED STATES IN 2021, THE GOVERNMENT IS SEEKING TO ANALYZE ACCIDENT DATA TO GATHER USEFUL INFORMATION FOR MAKING DECISIONS AND REGULATIONS AIMED AT REDUCING ACCIDENT RATES IN THE COUNTRY.

OBJECTIVE

TO DEVELOP EFFECTIVE STRATEGIES FOR REDUCING ROAD TRAFFIC ACCIDENT RATES IN 2022, I WILL ANALYZE TRAFFIC ACCIDENT DATA FROM 2021. BY EXAMINING THIS DATA, I AIM TO IDENTIFY PATTERNS AND FACTORS CONTRIBUTING TO ACCIDENTS, ENABLING ME TO PROPOSE TARGETED AND IMPACTFUL MEASURES FOR IMPROVING TRAFFIC SAFETY IN THE UPCOMING YEAR.

IDENTIFY FACTORS THAT CAUSE ACCIDENTS

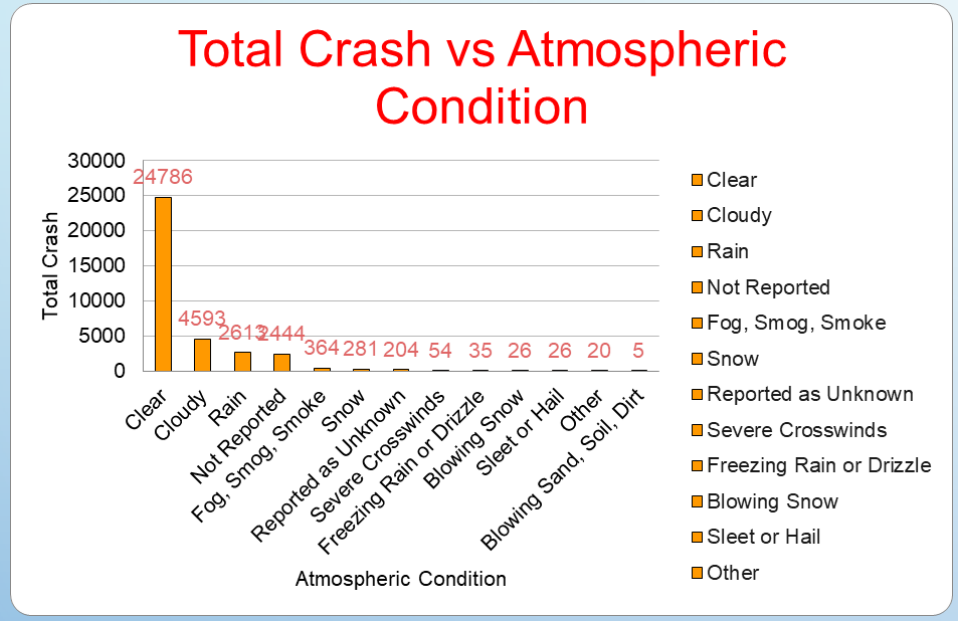
1. Identifying Conditions that Increase Accident Risk
2. Top 10 States with the Highest Number of Accidents
3. Average Number of Accidents Occurring Every Hour
4. Percentage of Accidents Caused by Drunk Drivers
5. Percentage of Accidents in Rural and Urban Areas
6. Number of Accidents by Day of the Week

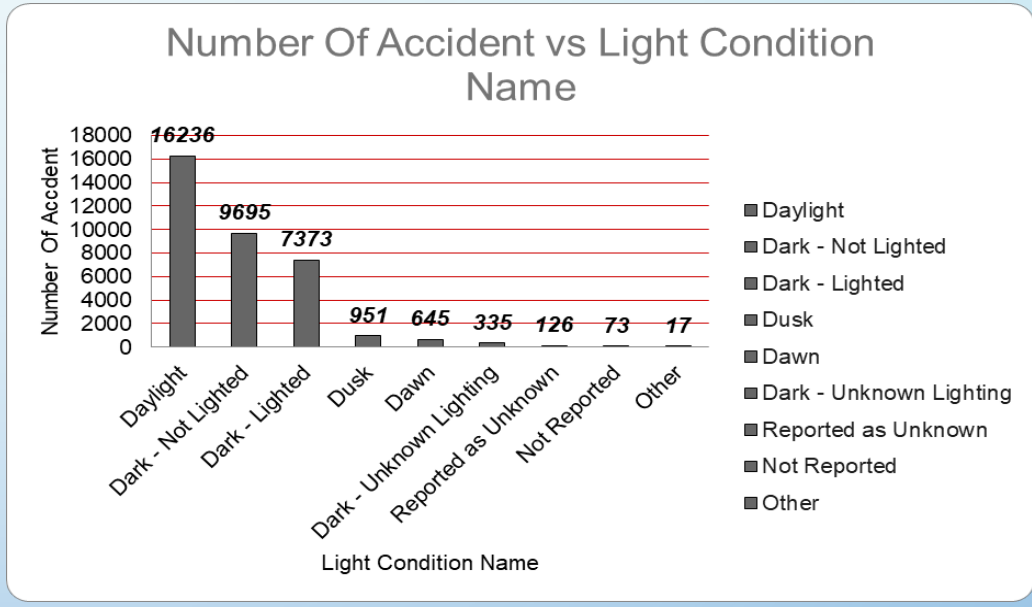
DATA CLEANING

- CREATE TABLE IF NOT EXISTS NEW_CRASH AS (WITH -- AWAL DARI CLEANSING CEK_DATA_KOSONG AS (SELECT *, CASE -- DATA KOSONG WHEN NUMBER_OF_PARKED_WORKING_VEHICLES IS NULL THEN 'DATA KENDARAAN PARKIR KOSONG' WHEN NUMBER_OF_FORMS_SUBMITTED_FOR_PERSONS_NOT_IN_MOTOR_VEHICLES IS NULL THEN 'JUMLAH KENDARAAN KOSONG' WHEN NUMBER_OF_PERSONS_IN_MOTOR_VEHICLES_IN_TRANSPORT_MVIT IS NULL THEN 'JUMLAH ORANG DALAM KENDARAAN KOSONG' WHEN NUMBER_OF_PERSONS_NOT_IN_MOTOR_VEHICLES_IN_TRANSPORT_MVIT IS NULL THEN 'JUMLAH ORANG DI LUAR KENDARAAN KOSONG' WHEN MILEPOINT IS NULL THEN 'MILSTONE KOSONG' WHEN NUMBER_OF_FATALITIES IS NULL THEN 'JUMLAH YANG FATAL KOSONG' WHEN NUMBER_OF_DRUNK_DRIVERS IS NULL THEN 'JUMLAH YANG MABUK KOSONG' WHEN TIMESTAMP_OF_CRASH IS NULL THEN 'WAKTU KOSONG' WHEN NUMBER_OF_VEHICLE_FORMS_SUBMITTED_ALL IS NULL THEN 'KENDARAAN KOSONG' WHEN NUMBER_OF_MOTOR_VEHICLES_IN_TRANSPORT_MVIT IS NULL THEN 'DATA KENDARAAN TABRAKAN KOSONG' WHEN STATE_NAME IS NULL THEN 'NAMA NEGARA BAGIAN KOSONG' WHEN LAND_USE_NAME IS NULL THEN 'NAMA KATEGORI DAERAH KOSONG' WHEN FUNCTIONAL_SYSTEM_NAME IS NULL THEN 'JENIS JALAN KOSONG' WHEN ATMOSPHERIC_CONDITIONS_1_NAME IS NULL THEN 'KONDISI CUACA KOSONG' WHEN MANNER_OF_COLLISION_NAME IS NULL THEN 'JENIS TABRAKAN KOSONG' WHEN TYPE_OF_INTERSECTION_NAME IS NULL THEN 'JENIS SIMPANG KOSONG' WHEN LIGHT_CONDITION_NAME IS NULL THEN 'KONDISI CAHAYA KOSONG' WHEN CITY_NAME IS NULL THEN 'NAMA KOTA KOSONG' -- DATA KENDARAAN YANG TERLIBAT 0 WHEN NUMBER_OF_VEHICLE_FORMS_SUBMITTED_ALL < 1 THEN 'KENDARAAN YANG TERLIBAT KECELAKAAN 0' -- DATA TIDAK SESUAI WHEN NUMBER_OF_VEHICLE_FORMS_SUBMITTED_ALL< NUMBER_OF_MOTOR_VEHICLES_IN_TRANSPORT_MVIT + NUMBER_OF_PARKED_WORKING_VEHICLES THEN 'DATA KENDARAAN YANG TERLIBAT TIDAK SESUAI' END VALIDASI FROM CRASH), -- AKHIR DARI DATA CLEANSING -- AWAL DARI KONVERSI WAKTU ZONAWAKTU AS (SELECT *, CASE -- +INTERVAL '-5 HOURS' WHEN STATE_NAME IN ('CONNECTICUT','DISTRICT OF COLUMBIA','DELAWARE','FLORIDA','GEORGIA','INDIANA','KENTUCKY','MAINE','MARYLAND','MASSACHUSETTS', 'MICHIGAN','NEW HAMPSHIRE','NEW JERSEY','NEW YORK','NORTH CAROLINA','OHIO','PENNSYLVANIA','RHODE ISLAND','SOUTH CAROLINA','TENNESSEE', 'VERMONT','VIRGINIA','WEST VIRGINIA') THEN TIMESTAMP_OF_CRASH AT TIME ZONE 'EST' -- + INTERVAL '-6 HOURS' WHEN STATE_NAME IN ('ALABAMA','ARKANSAS','FLORIDA','ILLINOIS','INDIANA','IOWA','KANSAS','KENTUCKY','LOUISIANA','MICHIGAN', 'MINNESOTA','MISSISSIPPI','MISSOURI','NEBRASKA','N.DAKOTA', 'OKLAHOMA','SOUTH DAKOTA','TENNESSEE','TEXAS','WISCONSIN') THEN TIMESTAMP_OF_CRASH AT TIME ZONE 'CST' --+ INTERVAL '-7 HOURS' WHEN STATE_NAME IN ('ARIZONA','ARIZONA','COLORADO','IDAHO','KANSAS','MONTANA','NEBRASKA','NEW MEXICO', 'NORTH DAKOTA','OREGON','SOUTH DAKOTA','TEXAS','UTAH','WYOMING') THEN TIMESTAMP_OF_CRASH AT TIME ZONE 'MST' -- + INTERVAL '-8 HOURS' ' WHEN STATE_NAME IN ('CALIFORNIA','IDAHO','NEVADA','OREGON','WASHINGTON') THEN TIMESTAMP_OF_CRASH AT TIME ZONE 'PST' --+ INTERVAL '-9 HOURS' WHEN STATE_NAME IN ('ALASKA') THEN TIMESTAMP_OF_CRASH AT TIME ZONE 'AKST' -- + INTERVAL '-10 HOURS' WHEN STATE_NAME IN ('HAWAII') THEN TIMESTAMP_OF_CRASH AT TIME ZONE 'HST' END WAKTULOKAL FROM CEK_DATA_KOSONG) -- AKHIR DARI KONVERFSI WAKTU SELECT * FROM ZONAWAKTU WHERE VALIDASI IS NULL);ALTER TABLE NEW_CRASH DROP VALIDASI;

IDENTIFYING CONDITIONS THAT INCREASE ACCIDENT RISK

ATMOSPHERIC CONDITION VS LIGHT CONDITION

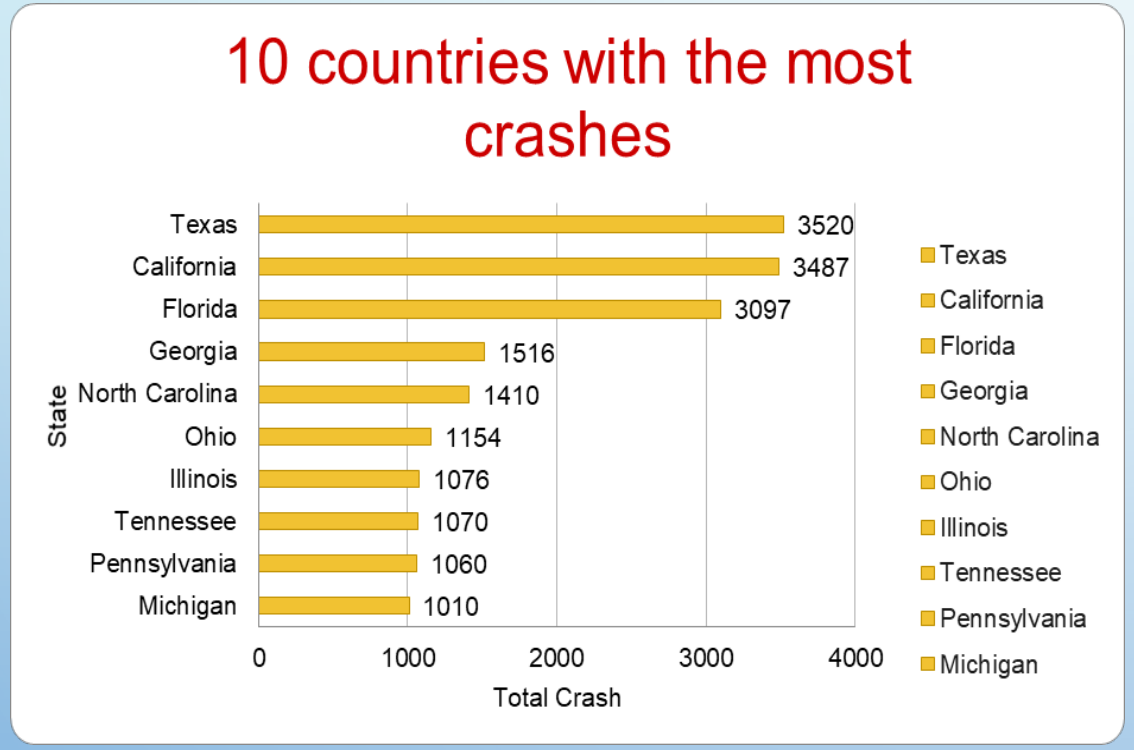




The majority of accidents occurred in clear weather (24,786), followed by cloudy (4,593) and rainy conditions (2,613).

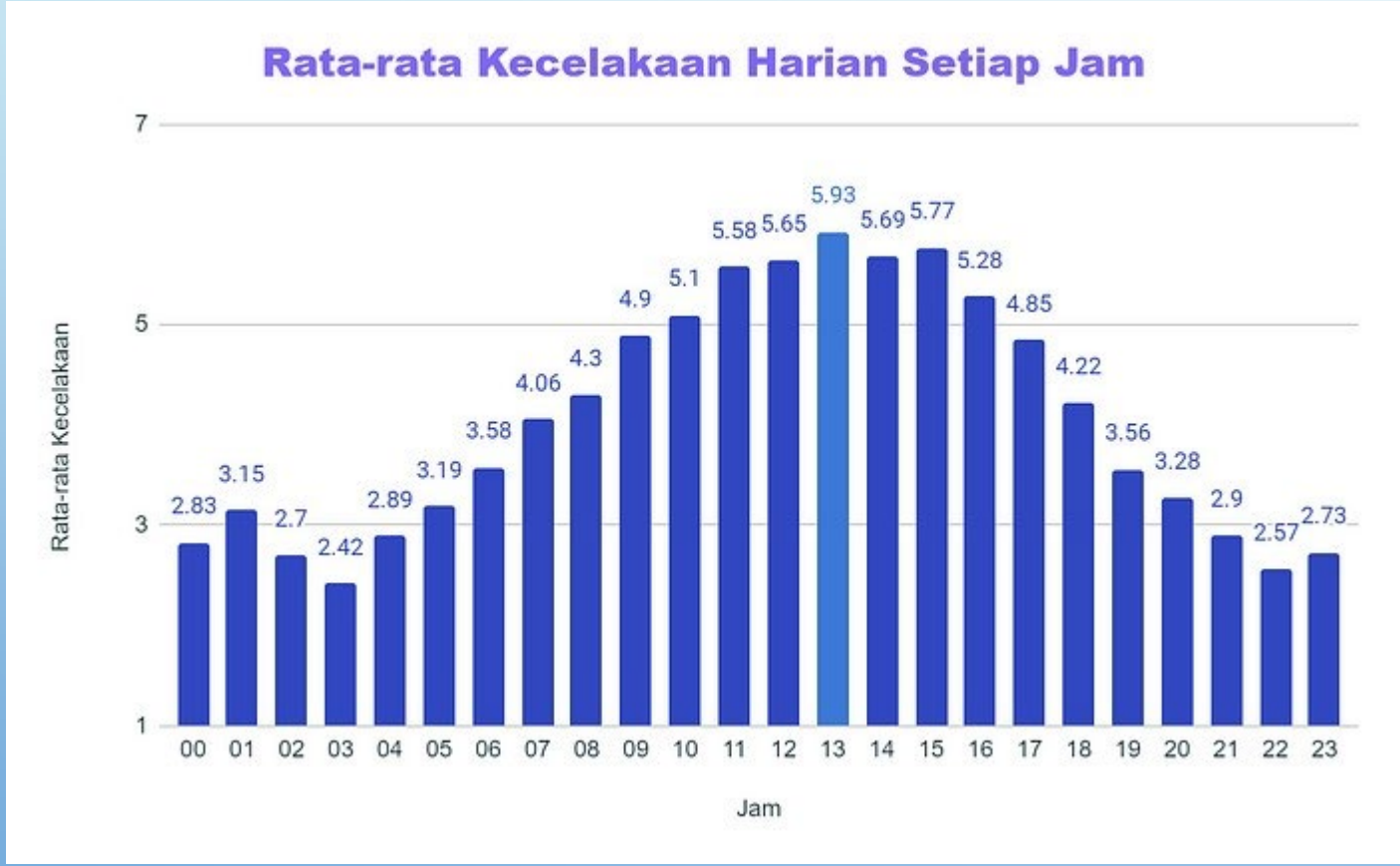
The majority of accidents occurred during daylight (16,236), followed by dark conditions where the area was not lighted (9,695) and lighted (7,373).
In bright daylight conditions, sunlight can glare and impair driver visibility. Clear weather often sees increased outdoor activity and vehicle use, leading to higher traffic density. This heightened traffic density can elevate accident risks, especially if drivers fail to maintain safe distances. Psychologically, drivers may feel more relaxed and secure in clear weather, potentially leading to negligence of traffic signs or safety precautions like driving at high speeds.

Top 10 States with the Highest Number of Accidents



The table shows the total number of crashes in various states. Texas leads with 3,520 crashes, followed by California with 3,487 and Florida with 3,097. Other states with notable crash numbers include Georgia (1,516), North Carolina (1,410), and Ohio (1,154). Illinois, Tennessee, Pennsylvania, and Michigan also report over 1,000 crashes each, with Illinois at 1,076, Tennessee at 1,070, Pennsylvania at 1,060, and Michigan at 1,010. The dense population in the state causes a high number of accidents in the state.

AVERAGE NUMBER OF ACCIDENTS PER DAY BASED ON THE TIME THE ACCIDENT OCCURRED

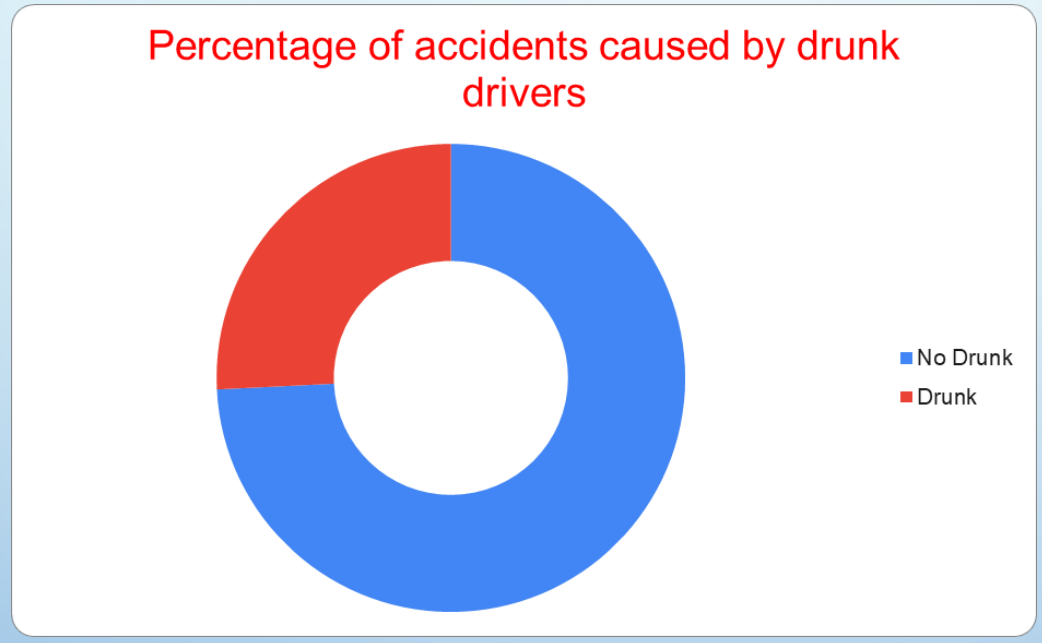


The highest average accident occurred at 13 o'clock

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

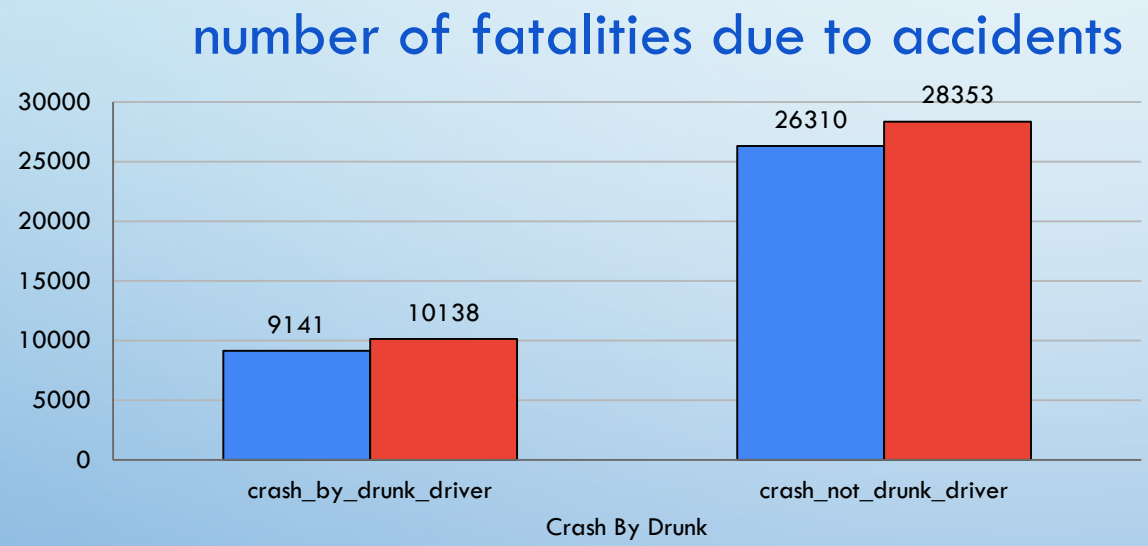
The lowest average accident occurs at 03 hours

Percentage of Accidents Caused by Drunk Drivers



Based on the comparison data of crashes caused by drunk drivers, we can conclude that 1 in 4 accidents are caused by drunk drivers. When we delve deeper into the fatality rates in these accidents, we find that the fatality rate in crashes caused by drunk drivers is higher than in those caused by sober drivers.

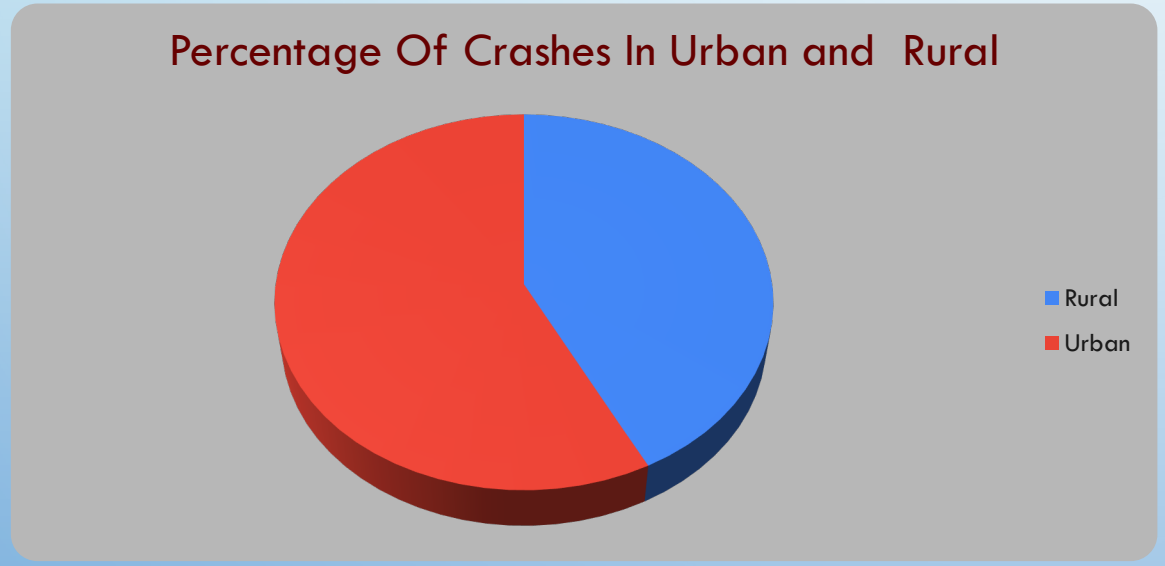
NUMBER OF FATALITIES DUE TO ACCIDENTS



There were 10138 fatalities from accidents caused by drunk drivers.

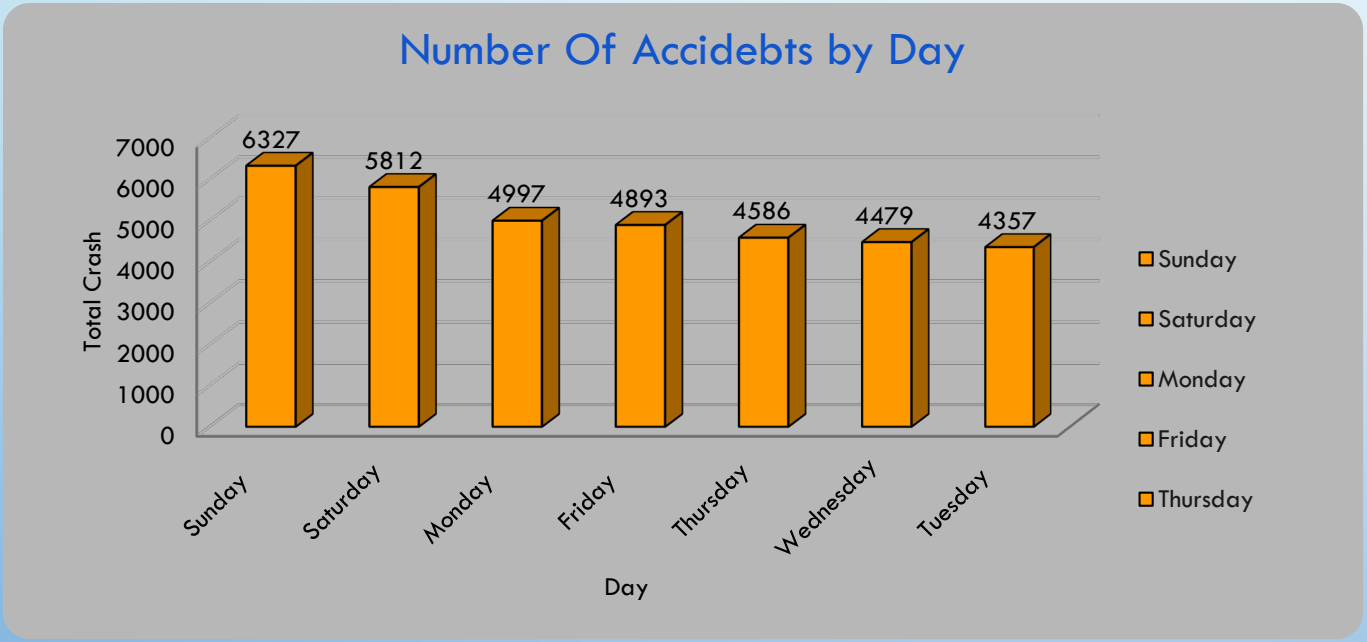
There were 28353 fatalities from accidents caused by non-drunk drivers.

ACCIDENTS IN RURAL AND URBAN AREAS



Total accidents in urban areas 20123 with a percentage of 57.5

NUMBER OF ACCIDENTS BY DAY



The highest number of accidents occurred on weekends, 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

CONLUSION

- ROM THE DATA ABOVE, WE CAN CONCLUDE THE FOLLOWING:
- **LIGHT CONDITIONS REVIEW**: REEVALUATE LIGHT CONDITION DEFINITIONS FOR MORE ACCURATE ANALYSIS.
- **DRIVER EDUCATION**: PROMOTE SAFE DRIVING HABITS DURING CLEAR WEATHER AND DAYLIGHT.
- **INFRASTRUCTURE LIGHTING**: IMPROVE LIGHTING ON POORLY LIT ROAD SECTIONS FOR BETTER VISIBILITY.
- **AVOID DUI (DRIVING UNDER THE INFLUENCE)**: RAISE AWARENESS ABOUT THE DANGERS OF DRIVING WHILE IMPAIRED.