
A COMPREHENSIVE DATA ANALYSIS OF MARYLAND STATEWIDE VEHICLE CRASHES

By

Varun Balle - QL24695

Sai Neelesh Kuntimala - YT93657

Dheeraj Kumar Boddu - DC32292

Meghna Aryasri – BH11498

Venkata Sridhar Perepu - SZ78217

PROBLEM STATEMENT:

As a data analyst, you have been hired and given the data of the vehicle crashes happening in Maryland, by the traffic police department in order to obtain analysis of the crashes which include high damage and alert the public about the certain scenarios and times, which might exercise more caution. The traffic department particularly wishes to know the counties where more crashes occur and the quarter in which they occur the most, while also analyzing the movement of the vehicle during the crash.

INTRODUCTION:

The human race has advanced significantly in the field of transportation as a whole, particularly in the last 50 years, from the creation of the wheel to modern automobiles. Even though we have taken many safeguards to develop our technologies, accidents nonetheless significantly impact how we live our daily lives. The unpredictable weather in Maryland has significantly changed transportation patterns. However, additional variables showed up in our findings, providing us with information about additional causes that may contribute to accidents in Maryland. Even if the making of an automobile had nothing to do with the weather, it nonetheless provided valuable insight into how automakers may be more ready for bad weather. Drivers and automakers must be ready for all weather conditions, especially given how unpredictable Maryland's weather may be, to reduce the likelihood of accidents. The state of Maryland uses crash reports, which aid in the analysis of accident data, to deal with crashes in a data- driven manner. The MDOT-State Highway Administration processes information from the statewide police crash report database (ACRS) to produce the crash data summary reports, which offer comprehensive aggregate data on motor vehicle crashes (general, injury, and fatal) as well as incidents at the jurisdictional level.

The main objective of our project "A comprehensive Data Analysis of Maryland Statewide Vehicle Crashes" is to analyze the data collected from the Maryland Crash Data Resources and extract meaningful insights such as nature and reasons for the accidents from the data. The Crash report data from January 2015 to June 2022 is collected by the Maryland State Police and we are using the 3 datasets such as Maryland Statewide Vehicle crashes, Maryland Vehicle details, Maryland statewide Vehicle crashes Person details, and performing the Exploratory Data Analysis on the data to get to know the detailed statistical Inference about accidents occurred in different counties of Maryland and analyze the Vehicle Crash data to get valuable insights by extracting meaningful patterns such as the occurrence of an accident and their relationship with other parameters in the data set.

DESCRIPTION OF THE DATASET:

The dataset we used is Maryland Crash Data Resources. The data for this dataset is provided by the Maryland state police. We used three datasets from this Resource page:

- Maryland_Statewide_Vehicle_Crashes
- Maryland_Statewide_Vehicle_Crashes_-_Person_Details_Anonymized_
- Maryland_Statewide_Vehicle_Crashes_-_Vehicle_Details

These datasets were merged on common columns and the final merged dataset was used for Exploratory data analysis.

DATASET 1 (MARYLAND_STATEWIDE_VEHICLE_CRASHES)

Column Name	Description	Type
YEAR	Year (Number) Comma Removed	Number
QUARTER	Quarter of the year in which the crash occurred	Plain Text
LIGHT_DESC	Describes the lighting conditions at the time of the crash	Plain Text
LIGHT_CODE	Numeric code associated with LIGHT_DESC	Number
COUNTY_DESC	Name of the county of the location of the crash	Plain Text
COUNTY_NO	Number corresponding to the COUNTY_DESC	Number
MUNI_DESC	Name of the municipality of the location of the crash	Plain Text
MUNI_CODE	Numeric code associated with the MUNI_DESC	Number
JUNCTION_DESC	Describes whether the crash happened at an intersection or a non-intersection	Plain Text
JUNCTION_CODE	Numeric code corresponding to the JUNCTION_DESC	Number
COLLISION_TYPE_DESC	Describes the nature of the collision, whether it is one-vehicle or head-on or same-direction etc.	Plain Text
COLLISION_TYPE_CODE	Numeric code corresponding to the COLLISION_TYPE_DESC	Number
SURF_COND_DESC	Describes the conditions of the road surface at the crash location	Plain Text
SURF_COND_CODE	Numeric code corresponding to the SURF_COND_DESC	Number

LANE_CODE	Numeric description of the number of lanes at the place of the crash	Number
RD_COND_DESC	Description of road conditions at the site of the crash whether there was a defect or no defect	Plain Text
RD_COND_CODE	Numeric code associated with RD_COND_DESC	Number
RD_DIV_DESC	Description of the Road Division type whether it was one-way or two-way	Plain Text
RD_DIV_CODE	Numeric code associated with RD_DIV_DESC	Number
FIX_OBJ_DESC	Details of the fixed object in case the collision involves a fixed object (curb, building, fence etc.,)	Plain Text
FIX_OBJ_CODE	Numeric code associated with FIX_OBJ_DESC	Number
REPORT_NO	Report number of the crash report	Plain Text
REPORT_TYPE	Type of report of the crash	Plain Text
WEATHER_DESC	Description of weather conditions at the time of the crash	Plain Text
WEATHER_CODE	Numeric code associated with WEATHER_DESC	Number
ACC_DATE	Date of occurrence of crash	Plain Text
ACC_TIME	Time of occurrence of the crash	Plain Text
LOC_CODE	Location code	Plain Text
SIGNAL_FLAG_DESC	Description of signal flag	Plain Text
SIGNAL_FLAG	Alphabetical code corresponding to SIGNAL_FLAG_DESC	Plain Text
C_M_ZONE_FLAG	Indication of a Control Manufacturing zone	Plain Text
AGENCY_CODE	Indicates the agency with an alphabetical code	Plain Text
AREA_CODE	Area code for the location of the crash	Plain Text
HARM_EVENT_DESC1	Details of the entities harmed during the crash-1	Plain Text
HARM_EVENT_CODE1	Numeric code associated with HARM_EVENT_DESC1	Number
HARM_EVENT_DESC2	Details of the entities harmed during the crash-2	Plain Text
HARM_EVENT_CODE2	Numeric code associated with HARM_EVENT_DESC2	Number
RTE_NO	Route Number	Number
ROUTE_TYPE_CODE	Type of route	Plain Text
RTE_SUFFIX	Route suffix if applicable	Plain Text

LOG_MILE	Log mile on the route closest to the location of the crash	Number
LOGMILE_DIR_FLAG_DESC	Log mile direction on the route (whether south or west etc.,)	Plain Text
LOGMILE_DIR_FLAG	Alphabetical code/flag corresponding to LOGMILE_DIR_FLAG_DESC	Plain Text
MAINROAD_NAME	Name of the main road where the crash occurred	Plain Text
DISTANCE	Distance	Number
FEET_MILES_FLAG_DESC	Describes whether the distance in the above is in miles or feet	Plain Text
FEET_MILES_FLAG	Alphabetical code/flag corresponding to FEET_MILES_FLAG_DESC	Plain Text
DISTANCE_DIR_FLAG	Alphabetical code/flag corresponding to direction (East, North etc.,)	Plain Text
REFERENCE_NO	Reference number	Number
REFERENCE_TYPE_CODE	Code corresponding to the reference type	Plain Text
REFERENCE_SUFFIX	Reference suffix	Plain Text
REFERENCE_ROAD_NAME	Reference road name	Plain Text
LATITUDE	The latitude coordinate of the location of the crash	Number
LONGITUDE	The longitude coordinate of the location of the crash	Number

DATASET 2 - MARYLAND_STATEWIDE_VEHICLE_CRASHES_-_PERSON_DETAILS
ANONYMIZED

Column Name	Description	Type
SEX_DESC	Sex of the person	Plain Text
SEX_CODE	Alphabetical code corresponding to the SEX_DESC	Plain Text

CONDITION_DESC	Condition of the person during driving (under influence or fatigued etc.,)	Plain Text
CONDITION_CODE	Numeric code associated with the CONDITION_DESC	Number
INJ_SEVER_DESC	Severity of the injury	Plain Text
INJ_SEVER_CODE	Numeric code associated with the INJ_SEVER_DESC	Number
REPORT_NO	Report number associated with the crash report	Plain Text
OCC_SEAT_POS_DESC	Crash occurrence seat position description	Plain Text
OCC_SEAT_POS_CODE	Numeric code associated with the OCC_SEAT_POS_DESC	Plain Text
PED_VISIBLE_DESC	Description of pedestrian if applicable	Plain Text
PED_VISIBLE_CODE	Numeric code associated with the PED_VISIBLE_DESC	Number
PED_OBEY_DESC	Description of whether the pedestrian obeyed the pedestrian signal	Plain Text
PED_OBEY_CODE	Numeric code associated with the PED_OBEY_DESC	Number
PED_TYPE_DESC	Pedestrian type (pedestrian, bicyclist etc.,)	Plain Text
PED_TYPE_CODE	Numeric code associated with PED_TYPE_DESC	Number
PED_LOCATION_CODE	Numeric code describing the location of the Pedestrian location	Number
MOVEMENT_DESC	Movement of the pedestrian involved during the crash	Plain Text
MOVEMENT_CODE	Numeric code associated with MOVEMENT_DESC	Number
PERSON_TYPE_DESC	Type of the person injured in the crash	Plain Text
PERSON_TYPE	Alphabetic representation of the PERSON_TYPE_DESC	Plain Text
ALCOHOL_TEST_DESC	Description of whether the alcohol test was given or refused	Plain Text

ALCOHOL_TEST_CODE	Numeric code associated with ALCOHOL_TEST_DESC	Number
ALCOHOL_TESTTYPE_DESC	Type of the alcohol test (blood, breath etc.,)	Plain Text
ALCOHOL_TESTTYPE_CODE	Numeric code associated with ALCOHOL_TESTTYPE_DESC	Number
DRUG_TEST_DESC	Description of whether the drug test was given or refused	Plain Text
DRUG_TEST_CODE	Numeric code associated with DRUG_TEST_DESC	Number
DRUG_TESTRESULT_DESC	Results of the Drug test	Plain Text
DRUG_TESTRESULT_CODE	Numeric code associated with DRUG_TESTRESULT_DESC	Plain Text
BAC_CODE	Numeric code associated with Blood alcohol concentration	Number
FAULT_FLAG_DESC	Fault flag description	Checkbox
FAULT_FLAG	Alphabetic code associated with FAULT_FLAG_DESC	Plain Text
EQUIP_PROB_DESC	Equipment problem description	Plain Text
SAF_EQUIP_DESC	Description of safety equipment used	Plain Text
SAF_EQUIP_CODE	Numeric code associated with SAF_EQUIP_DESC	Number
EJECT_DESC	Description of whether ejected or not	Plain Text
EJECT_CODE	Numeric code associated with EJECT_DESC	Number
DATE_OF_BIRTH	Date of birth of person (inconsistent format)	Plain Text
PERSON_ID	Person ID	Plain Text
LICENSE_STATE_CODE	Two alphabet code representing the state of the License plate	Plain Text
CLASS	Class of the Driver's License of the person driving	Plain Text
CDL_FLAG_DESC	Whether or not it is CDL class	Checkbox

CDL_FLAG	Alphabetical associated with CDL_FLAG_DESC	Checkbox
VEHICLE_ID	Unique vehicle ID	Plain Text
EMS_UNIT_LABEL	EMS unit label	Plain Text
AIRBAG_DEPLOYED	Numeric code indicating whether airbag was deployed	Plain Text
YEAR	Year of the accident	Plain Text
Quarter	Quarter of the year in which the crash occurred	Plain Text

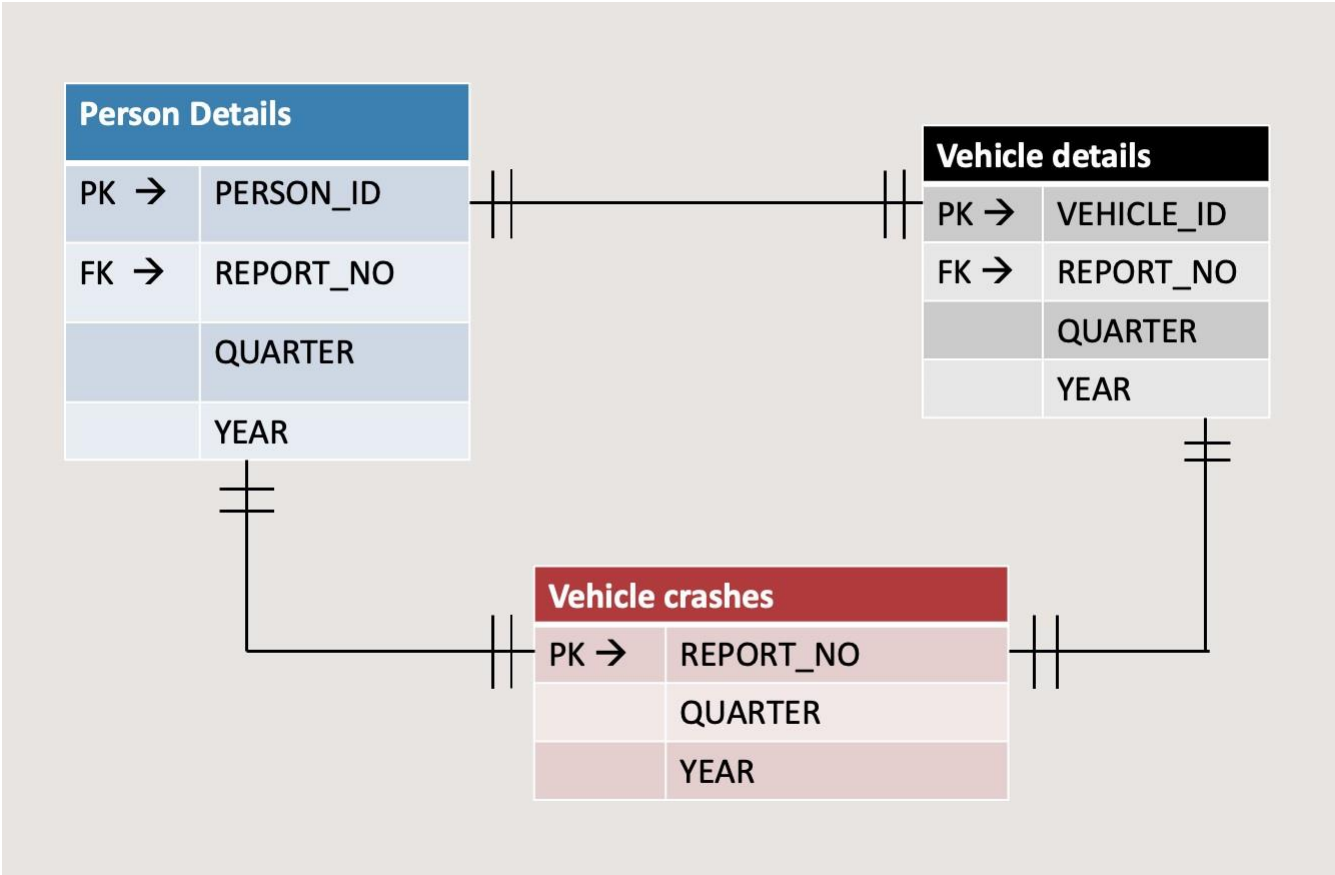
DATASET 3 - MARYLAND_STATEWIDE_VEHICLE_CRASHES_-_VEHICLE_DETAILS

Column Name	Description	Type
HARM_EVENT_DESC	Entity harmed in the crash	Plain Text
HARM_EVENT_CODE	Numeric code associated with HARM_EVENT_DESC	Number
CONTI_DIRECTION_DESC	Description of the direction of the vehicle	Plain Text
CONTI_DIRECTION_CODE	Numeric code associated with CONTI_DIRECTION_DESC	Plain Text
DAMAGE_DESC	Description of the severity of the crash	Plain Text
DAMAGE_CODE	Numeric code associated with DAMAGE_DESC	Number
MOVEMENT_DESC	Description of movement of the affected vehicle	Plain Text
MOVEMENT_CODE	Numeric code associated with MOVEMENT_DESC	Number
VIN_NO	License plate number	Plain Text
REPORT_NO	Unique report number	Plain Text

CV_BODY_TYPE_DESC	Type of body type	Plain Text
CV_BODY_TYPE_CODE	Numeric code associate CV_BODY_TYPE_DESC	Number
VEH_YEAR	Year of Vehicle manufacturing	Number
VEH_MAKE	Vehicle make	Plain Text
VEH_MODEL	Vehicle model	Plain Text
COMMERCIAL_FLAG_DESC	Description of whether commercial or not	Plain Text
COMMERCIAL_FLAG	Alphabetical code associated with COMMERCIAL_FLAG_DESC	Plain Text
HZM_NUM	HZM number	Plain Text
TOWED_AWAY_FLAG_DESC	Describes whether the vehicle was towed away	Plain Text
TOWED_AWAY_FLAG	Alphabetical code associated with TOWED_AWAY_FLAG_DESC	Plain Text
NUM_AXLES	Number of axles	Number
GVW_DESC	Gross weight description	Plain Text
GVW_CODE	Numeric code associated with GVW_DESC	Number
HIT_AND_RUN_FLAG_DESC	Whether it was a hit and run	Plain Text
HIT_AND_RUN_FLAG	Alphabetical code associated with HIT_AND_RUN_FLAG_DESC	Plain Text
HAZMAT_SPILL_FLAG_DESC	Yes/No description for Hazmat spill	Plain Text
HAZMAT_SPILL_FLAG	Alphabetical code associated with HAZMAT_SPILL_FLAG_DESC	Plain Text
VEHICLE_ID	Vehicle ID	Plain Text
TOWED_VEHICLE_CONFIG_D	Towed Vehicle Description	Plain Text

ESC		
TOWED_VEHICLE_CONFIG_CODE	Towed Vehicle configuration code	Number
AREA_DAMAGED_CODE_IMP1	Area Damaged code	Number
AREA_DAMAGED_CODE1	Area Damaged code 1	Number
AREA_DAMAGED_CODE2	Area Damaged code 2	Number
AREA_DAMAGED_CODE3	Area Damaged code 3	Number
AREA_DAMAGED_CODE_MAIN_D ESC	Area Damaged code main part	Plain Text
AREA_DAMAGED_CODE_MAIN	Numeric code associated with	Number

ER DIAGRAM:



SQL TO ANALYZE THE DATA:

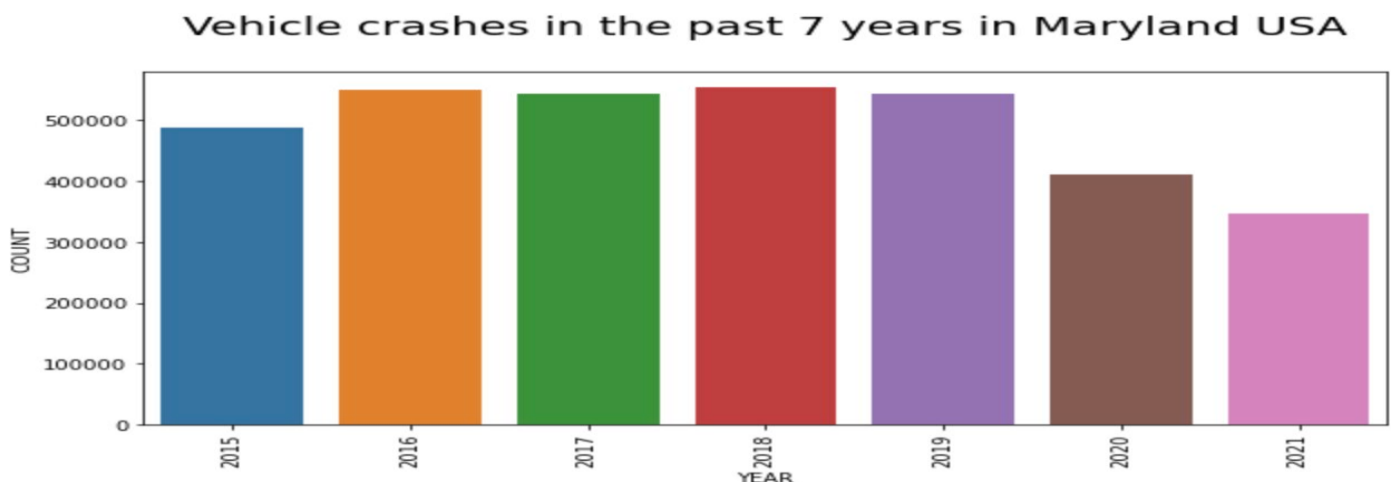
The data relating to vehicle crashes are analyzed based on the different factors that might have affected the occurrence of crashes. Various queries perform joins on different tables to analyze the data.

THE NUMBER OF CRASHES IN EACH YEAR:

```
SELECT PD.YEAR,COUNT(*) FROM Vehicle_Details VD JOIN Person_Details PD ON  
VD.REPORT_NO = PD.REPORT_NO AND
```

```
VD.YEAR = PD.YEAR AND VD.Quarter = PD.Quarter JOIN Vehicle_Crashes VC ON  
VD.REPORT_NO = VC.REPORT_NO AND
```

```
VD.YEAR = VC.YEAR AND VD.Quarter = VC.Quarter GROUP BY PD.YEAR
```

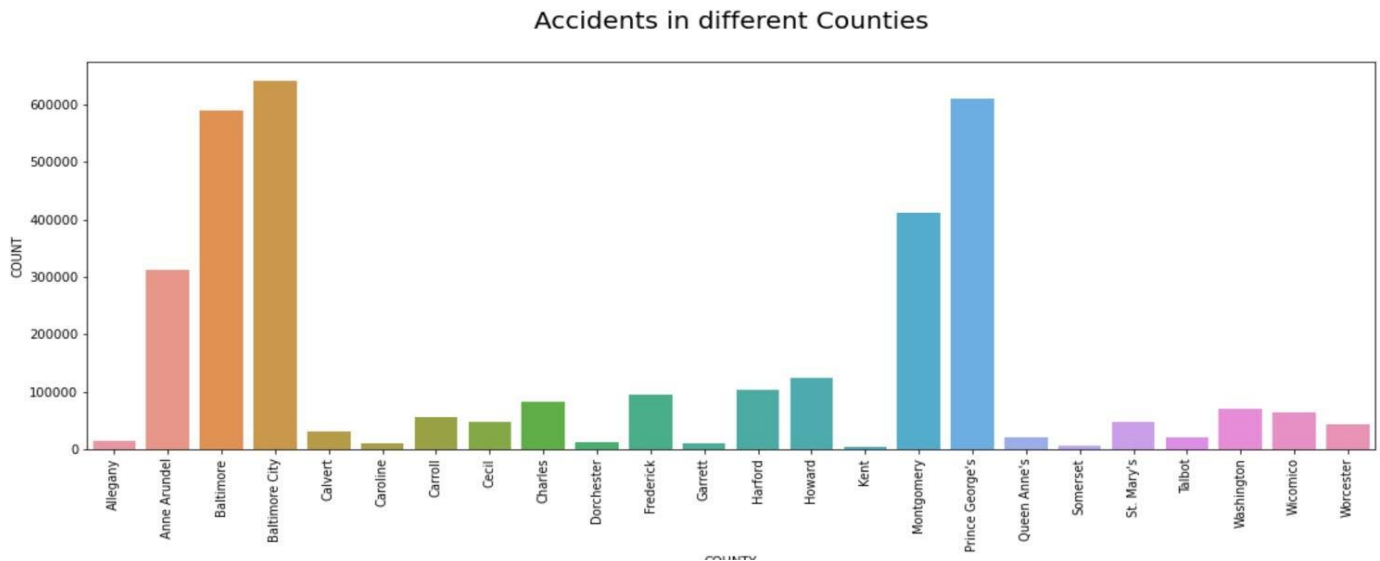


CRASHES IN DIFFERENT COUNTIES:

```
SELECT VC.COUNTY_DESC,COUNT(*) FROM Vehicle_Details VD JOIN  
Person_Details PD ON VD.REPORT_NO = PD.REPORT_NO AND
```

```
VD.YEAR = PD.YEAR AND VD.Quarter = PD.Quarter JOIN Vehicle_Crashes VC ON  
VD.REPORT_NO = VC.REPORT_NO AND
```

VD.YEAR = VC.YEAR AND VD.Quarter = VC.Quarter GROUP BY
VC.COUNTY_DESC

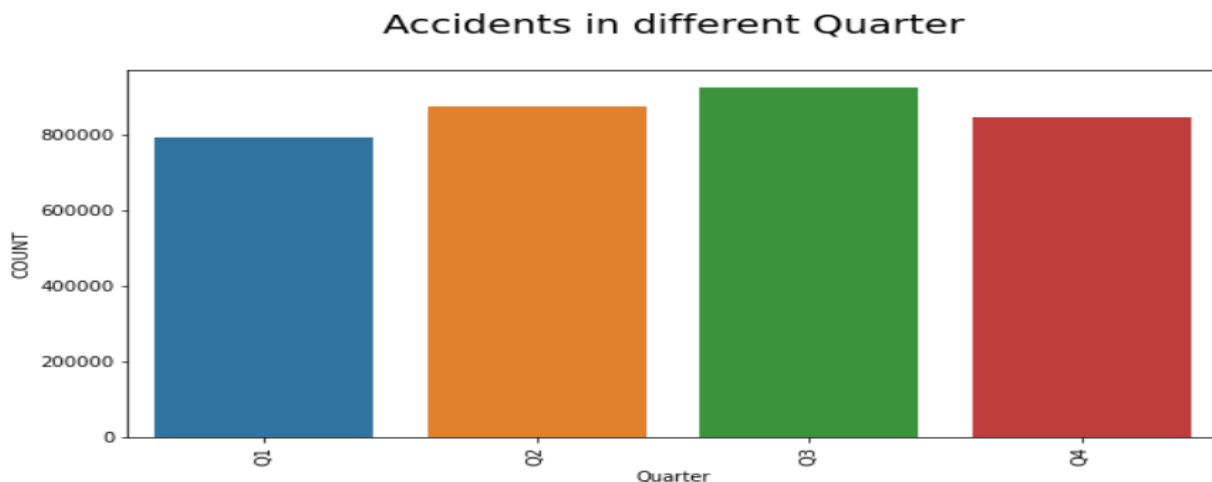


CRASHES INDIFFERENT QUARTERS

SELECT VC.Quarter, COUNT(*) FROM Vehicle_Details VD JOIN Person_Details PD ON
VD.REPORT_NO = PD.REPORT_NO AND

VD.YEAR = PD.YEAR AND VD.Quarter = PD.Quarter JOIN Vehicle_Crashes VC ON
VD.REPORT_NO = VC.REPORT_NO AND

VD.YEAR = VC.YEAR AND VD.Quarter = VC.Quarter GROUP BY VC.Quarter



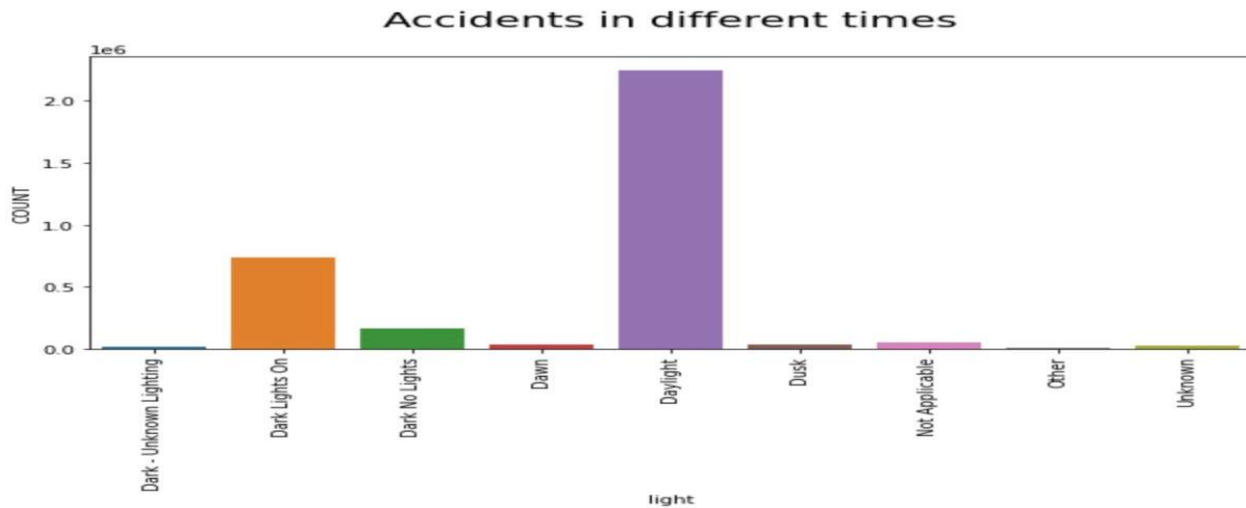
CRASH AT DIFFERENT TIMES

SELECT VC.LIGHT_DESC,COUNT(*) FROM Vehicle_Details VD
JOIN Person_Details PD ON VD.REPORT_NO =

PD.REPORT_NO AND

VD.YEAR = PD.YEAR AND VD.Quarter = PD.Quarter JOIN Vehicle_Crashes VC ON
VD.REPORT_NO = VC.REPORT_NO AND

VD.YEAR = VC.YEAR AND VD.Quarter = VC.Quarter GROUP BY



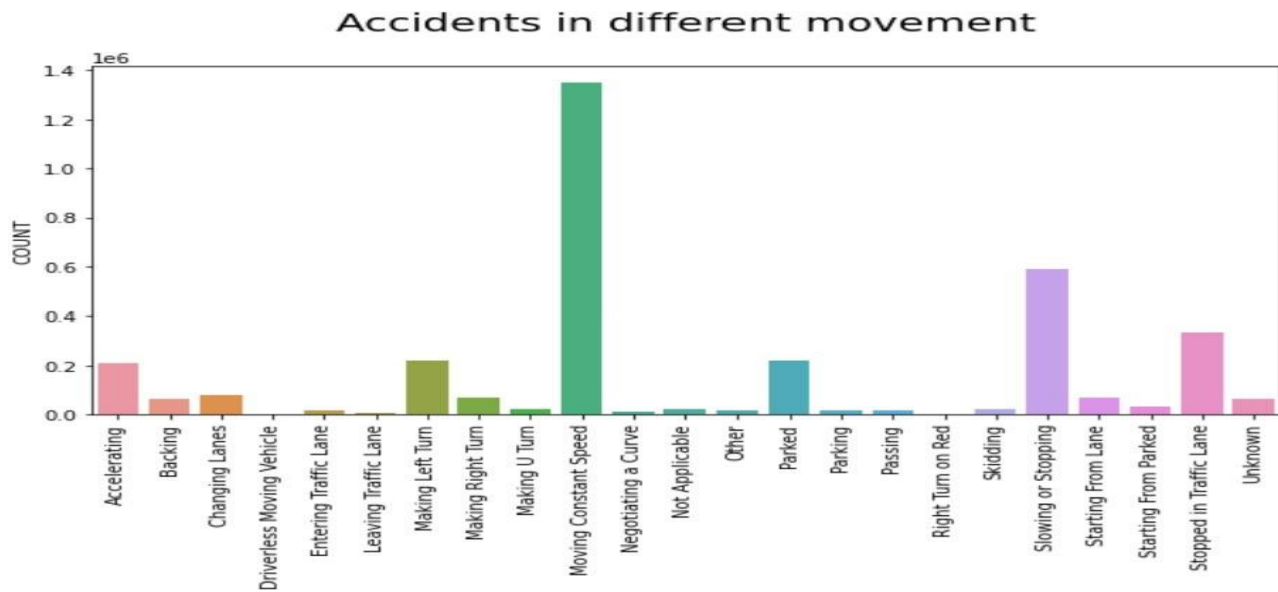
VC.LIGHT_DESC

MOVEMENT OF THE VEHICLE DURING THE CRASH:

SELECT VD.MOVEMENT_DESC,COUNT(*) FROM Vehicle_Details VD JOIN
Person_Details PD ON VD.REPORT_NO = PD.REPORT_NO AND

VD.YEAR = PD.YEAR AND VD.Quarter = PD.Quarter JOIN Vehicle_Crashes VC ON
VD.REPORT_NO = VC.REPORT_NO AND

VD.YEAR = VC.YEAR AND VD.Quarter = VC.Quarter GROUP BY
VD.MOVEMENT_DESC

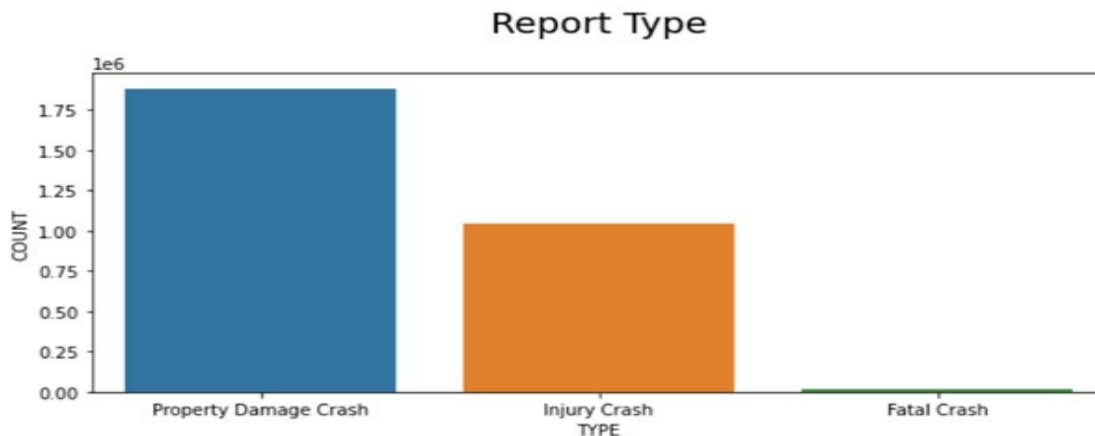


DAMAGE CAUSED IN THE CRASHES BASED ON REPORT DETAILS:

```
SELECT VC.REPORT_TYPE,COUNT(*) FROM Vehicle_Details VD JOIN
Person_Details PD ON VD.REPORT_NO = PD.REPORT_NO AND
```

```
VD.YEAR = PD.YEAR AND VD.Quarter = PD.Quarter JOIN Vehicle_Crashes VC ON
VD.REPORT_NO = VC.REPORT_NO AND
```

```
VD.YEAR = VC.YEAR AND VD.Quarter = VC.Quarter GROUP BY
VC.REPORT_TYPE
```



RESULTS:

- Maryland's dominant weather under crashes is rain followed by cloudy weather

- A drop in the number of crashes in 2020-21 due to the pandemic
- Baltimore County has the greatest number of crashes followed by Prince George
- Possible influence from summer break in Q3.
- Most of the crashes resulted in Property damage, while fatal crashes were low.
- Many crashes took place during daylight.

Discussion and future work:

From the detailed analysis of Maryland Statewide crashes, we found certain observations to be interesting. The interesting thing we found is that the accidents are dependent on other attributes such as weather, gender, type of vehicle and proficiency level of the driver, etc. From the data, the most intriguing observation we found was that most crashes occur at the non-Intersection junctions compared to the Intersection junctions. Another major work that could be done is adding certain attributes to perform regression and classification on the given data. This information about the analysis of Maryland statewide crashes and detailed conclusions about the occurrence of the accidents can be used by insurance companies to frame vehicle insurance policies

REFERENCES:

- *Maryland Crash and Traffic Fatalities Data*. (n.d.). Zero Deaths Maryland & Vision Zero - Maryland Highway Safety Office. <https://zerodeathsmd.gov/resources/crashdata/>
- *Maryland Statewide Vehicle Crashes | Open Data | opendata.maryland.gov*. (n.d.). Opendata.maryland.gov. [https://opendata.maryland.gov/](https://opendata.maryland.gov/Public-Safety/Maryland-Statewide-Vehicle-Crashes/65du-s3qu) Public-Safety/Maryland-Statewide-Vehicle- Crashes/65du-s3qu
- Willison, S. (2022, December 13). *csvs-to-sqlite*. GitHub. <https://github.com/simonw/csvs-to-sqlite>
- Database Design – 2nd Edition <https://opentextbc.ca/dbdesign01/>