NHTSA
Collision Dataset Analysis

Project Engineer: Dheeraj Kallakuri

Project Sponsor: Brendan Russo, NAU

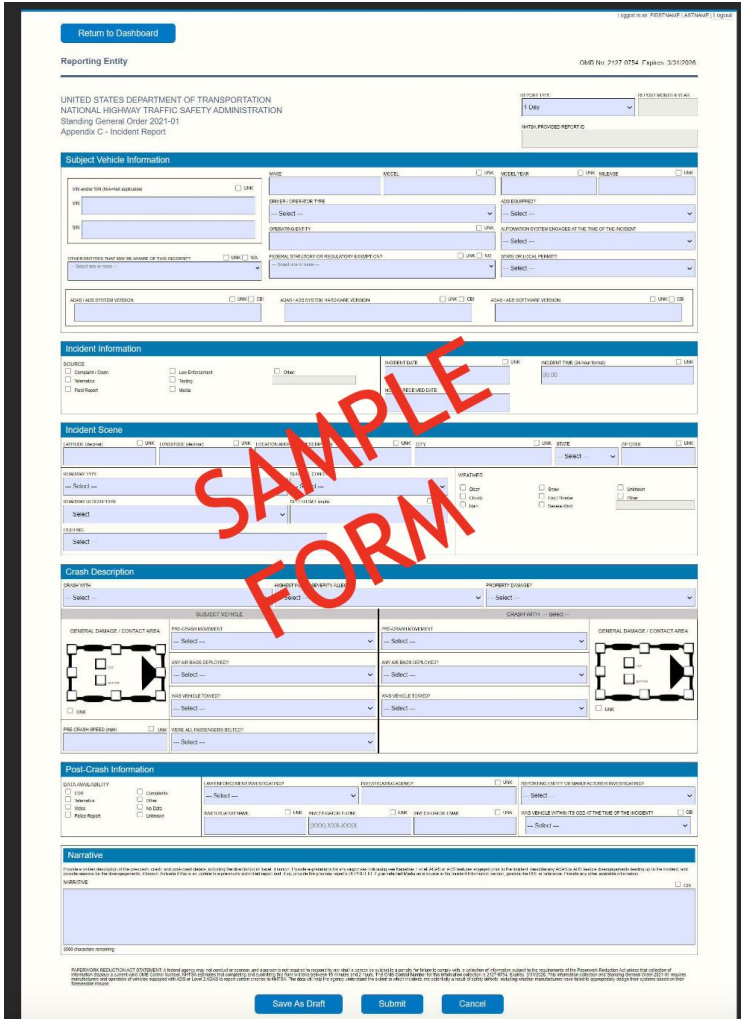
Overview

- Scope Summary
- Data Summary
- Data Analysis
- Data Visualization
- Automated vs Human driven vehicles.
- Predictive Modeling
- Future Work
- Takeaways
- References
- Q/A

Scope Summary

- **General Order**: The NHTSA has mandated that certain manufacturers and operators must report specific crashes involving vehicles equipped with ADS and ADAS Level 2 vehicles. The new version was amended in April 2023
- ADS: Automated Driving Systems(ADS)
- ADAS Level 2: Advanced Driver Assistance Systems(ADAS)
- Study the data and analyse the collision dataset and summarize the required data analysis and trends.
- Compare these collision elements with Human-driven vehicles collision elements.

ADS-ADAS Level 2 Incident Report Sample Form



Data Summary

- ADS has about 703 and ADAS Level 2 has 1154 data points and there are 137 columns. It has the information **from 2021 till 15 Sep 2023.**
- The following data comprises of **all the accident reported in US.**
- Human driven vehicle data of Arizona was been provided by sponsor and the data set is huge it has information **from 2010 to 2022.**
- It was tedious to load and data process back till 2010 so used 2020-2022 data facts **provided by ADOT.**
- Each year there there are 100k road accidents happening Arizona. There are more than 50 inputs. The following data holds the drivers involved, units and incident data.
- All Data are crucial for accurate interpretation and analysis
- Precise location data, CBI related data are all be redacted from the csv.

Steps

- Data Cleaning and Preprocessing
- Exploratory Data Analysis (EDA)
- Correlation between variable and factors
- Data visualization between columns: bar, pie, grouped bar graph
- Statistical analysis
- Descriptive statistics
- Natural language analysis on the narrative of the accident
- Prediction of Crashing partner?

Data Analysis

1. Around 42 analysis were done with ADS and ADAS Level 2 Data.
    a. Geospatial analysis: City wise, State wise, etc.
    b. Factor Analysis: Road Type, Weather type, Speed Wise, etc.
    c. Crash analysis: Injury Severity, Contact area, pre-crash movement etc.
    d. Reporting analysis, Safety analysis etc.
2. Around 10 comparison analysis between ADS/ADAS Level 2 data with human driven vehicles.
    a. Factor analysis
    b. Crash with analysis
3. Created a model using machine learning algorithm on the provided ADS and ADAS Level 2 data to with which these vehicles might crash is **predicted.**

Data Visualization of count of accidents

- Which?
- Who?
- Where?
- When?
- What?
- Why?

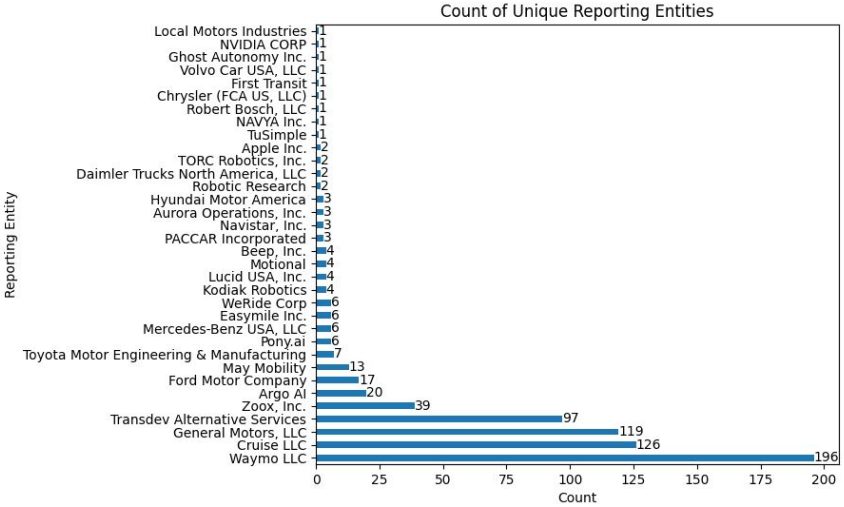
Data Visualization

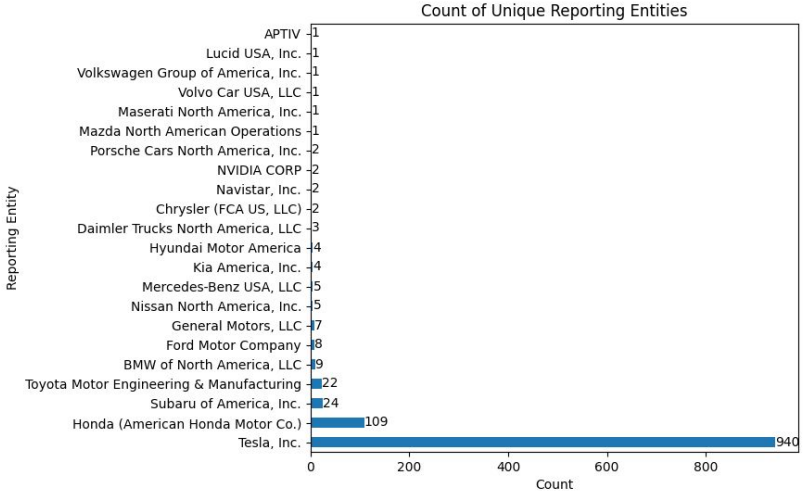
- Which?: reporting companies count
- Who?: Operator analysis
- Where?: city ,state
- When?: Incident time
- What?: weather, road type, mileage, speed limit
- Why?: precrash movements

Data Visualization

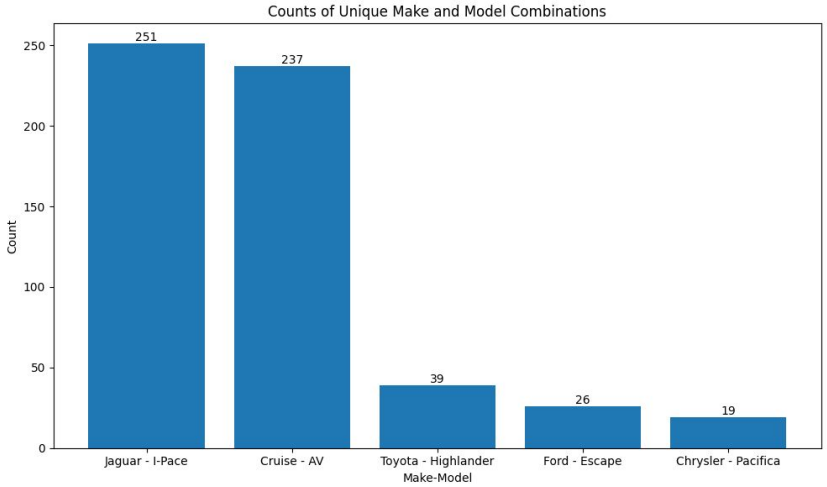
- **Which?: reporting companies count**

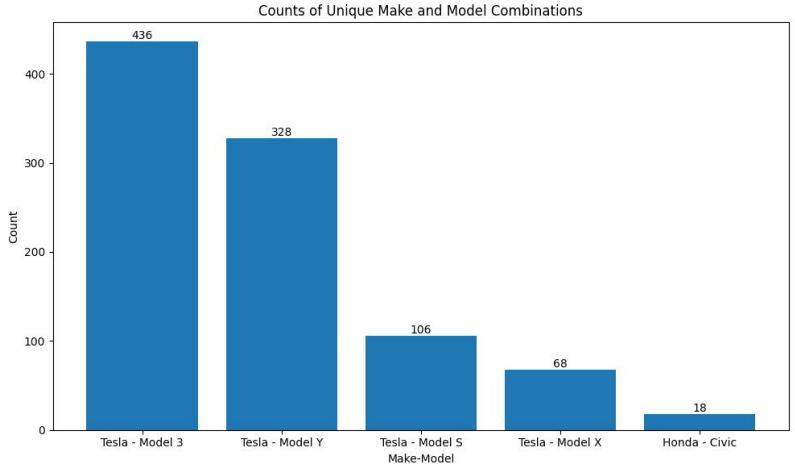
ADS, ADAS Level 2 Entity analysis



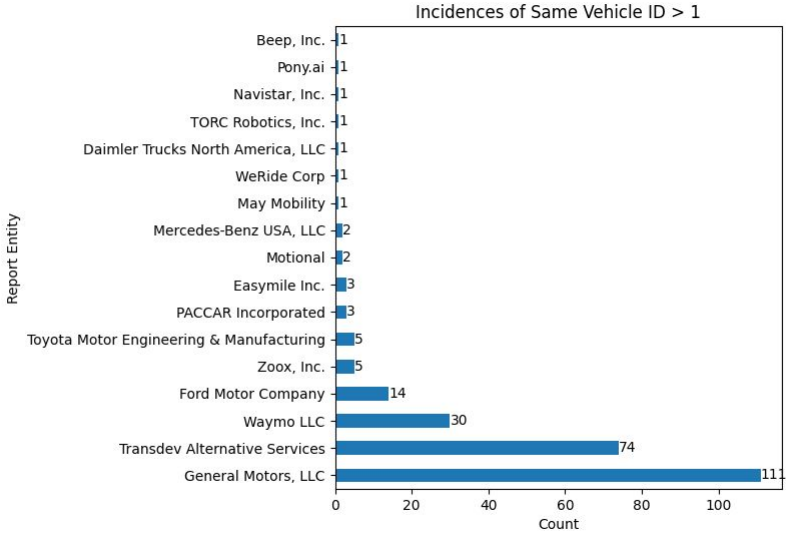


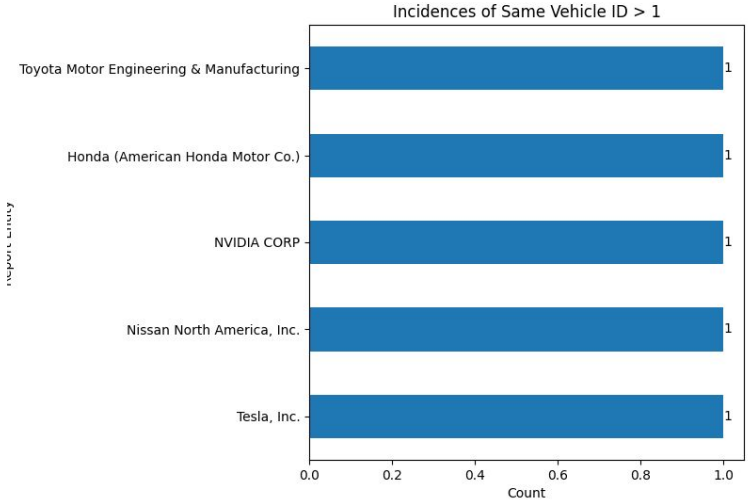
ADS, ADAS Level 2 Make-Model analysis >15





ADS, ADAS Level 2 Same Vehicle involved in crash

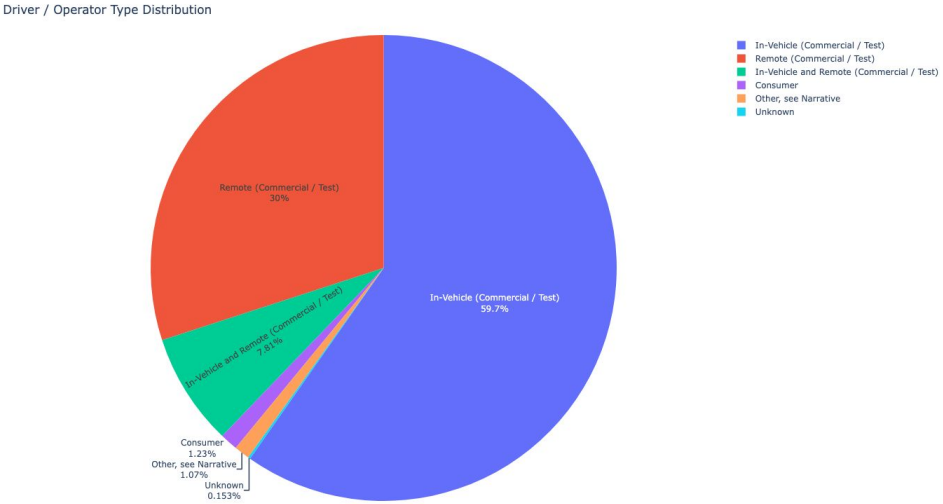


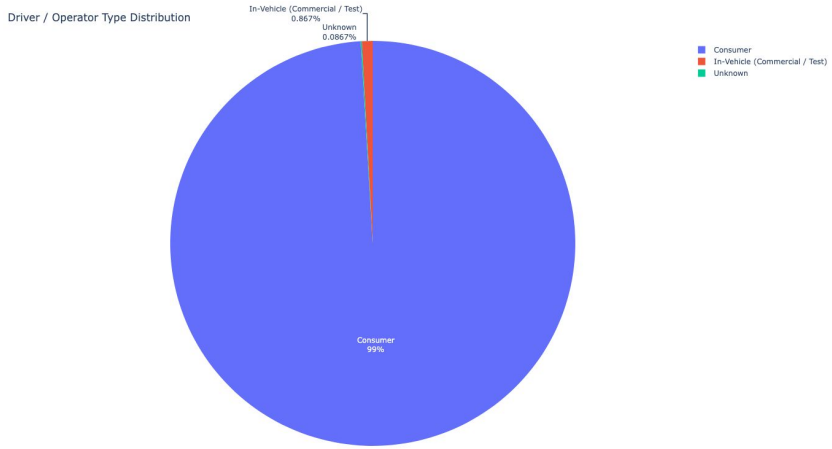


Data Visualization

- Which?: reporting companies count
- **Who?: Operator analysis**

ADS, ADAS Level 2 Driver/Operating type





Data Visualization

- Which?: reporting companies count
- Who?: Operator analysis
- **Where?: city ,state**

ADS, ADAS Level 2 City analysis

ads_city_analysis

City	count
San Francisco	420
Austin	37
Phoenix	36
Tempe	35
Las Vegas	13
Miami	12
Los Angeles	10
Chandler	9
Mesa	6
Santa Monica	5

adas_city_analysis

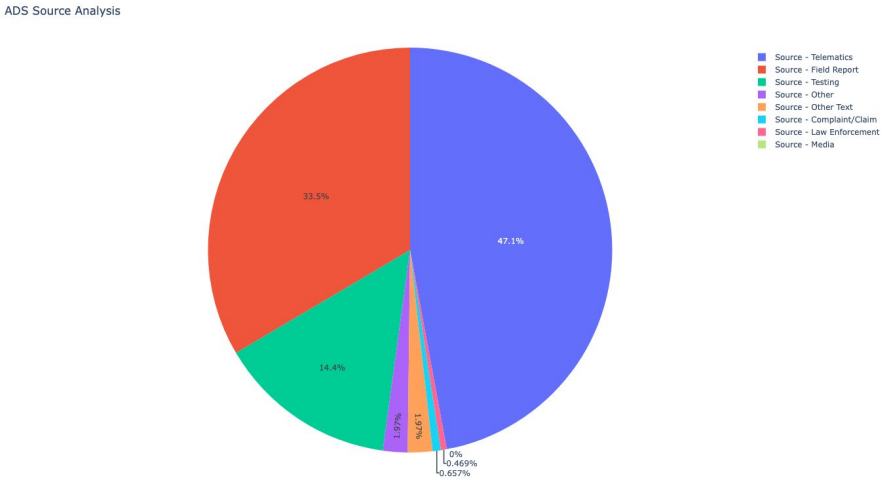
City	count
Los Angeles	42
Houston	15
San Jose	13
Dallas	11
Mountain View	11
San Diego	11
Fremont	10
Phoenix	10
San Antonio	9
Miami	8

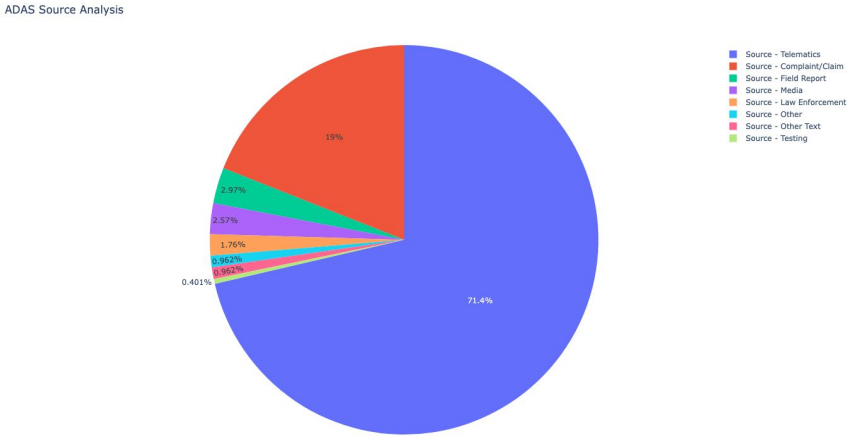
ADS, ADAS Level 2 State analysis

State	count
CA	469
AZ	96
TX	60
FL	23
NV	16
DC	6
MI	6
MN	5
NM	4
CO	4

State	count
CA	377
TX	101
FL	92
NY	52
NJ	42
GA	35
PA	33
VA	32
WA	28

ADS, ADAS Level 2 Source Analysis

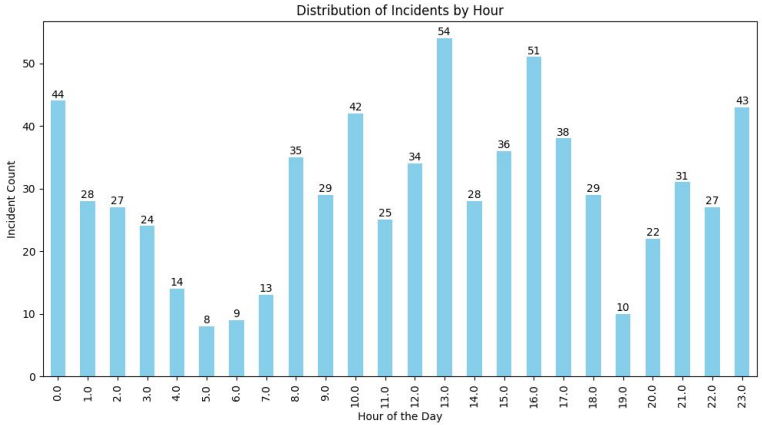


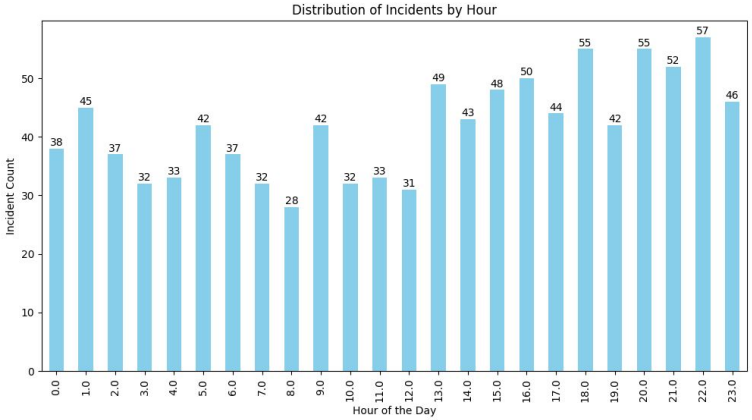


Data Visualization

- Which?: reporting companies count
- Who?: Operator analysis
- Where?: city ,state
- **When?: Incident time**

ADS, ADAS Level 2 Incident Time Analysis

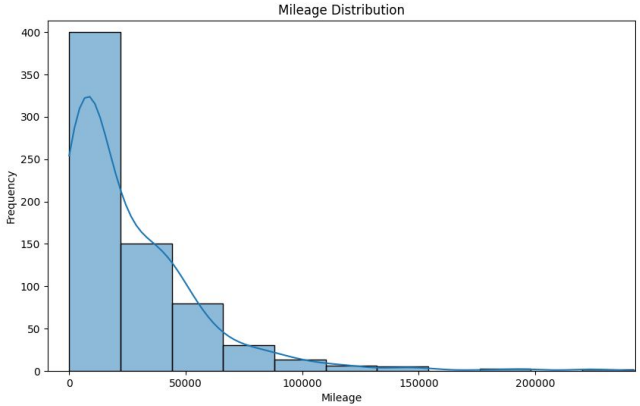




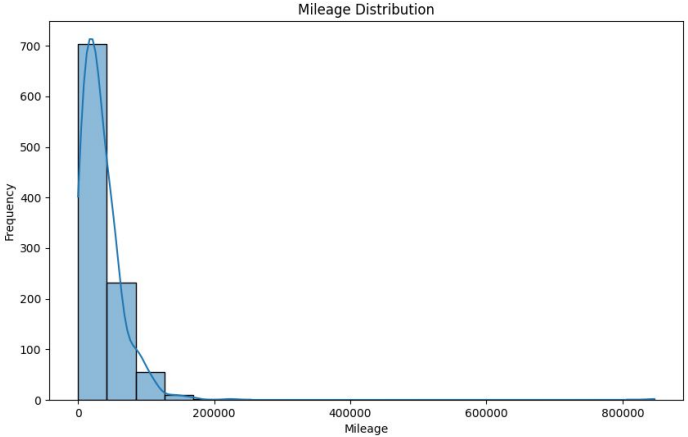
Data Visualization

- Which?: reporting companies count
- Who?: Operator analysis
- Where?: city ,state
- When?: Incident time
- **What?: weather, road type, mileage, speed limit**

ADS, ADAS Level 2 Mileage

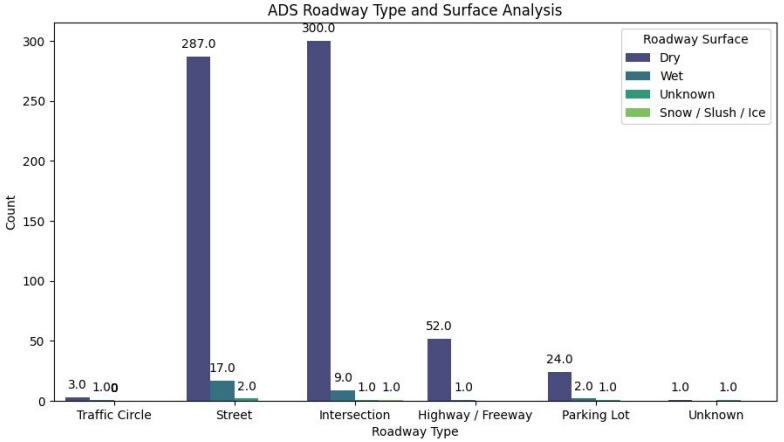


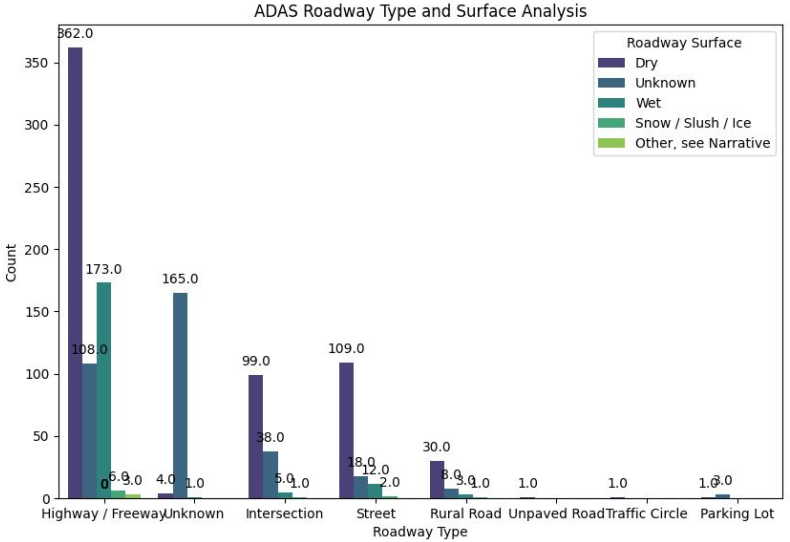

```
Summary Statistics:
count       690.000000
mean      27314.134783
std       34242.628596
min           0.000000
25%        5582.000000
50%       16827.000000
75%       39252.000000
max      440273.000000
Name: Mileage, dtype: float64

Median Mileage: 16827.0
Mileage Variance: 1172557613.1791947
Mileage Standard Deviation: 34242.628596227754
```



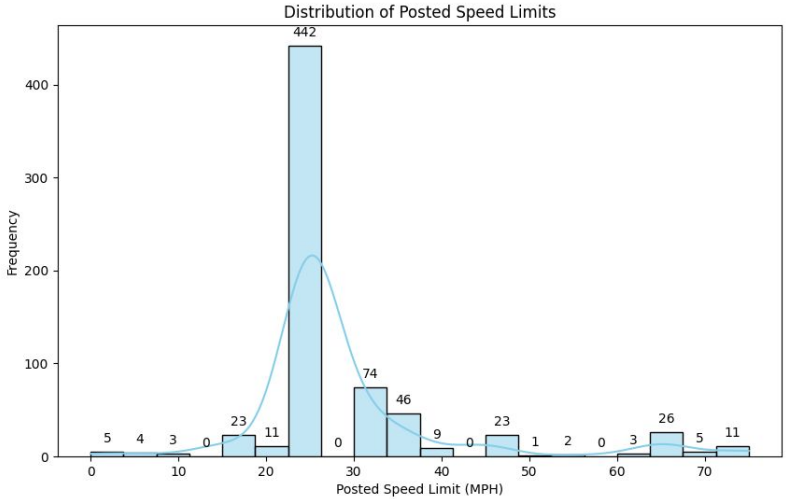

```
Summary Statistics:
count      1002.000000
mean      34642.550898
std       38146.140195
min          48.000000
25%       13557.000000
50%       27073.500000
75%       47641.500000
max      846777.000000
Name: Mileage, dtype: float64

Median Mileage: 27073.5
Mileage Variance: 1455128011.7501538
Mileage Standard Deviation: 38146.140194653424
```

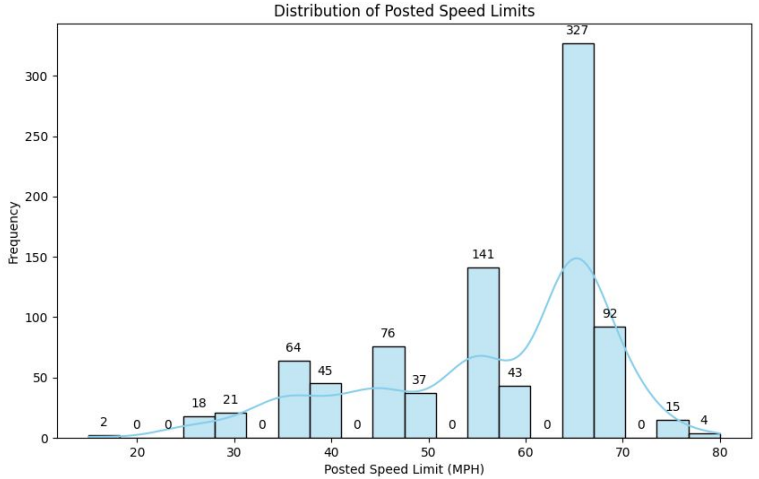
ADS, ADAS Level 2 Road Type/Surface Analysis



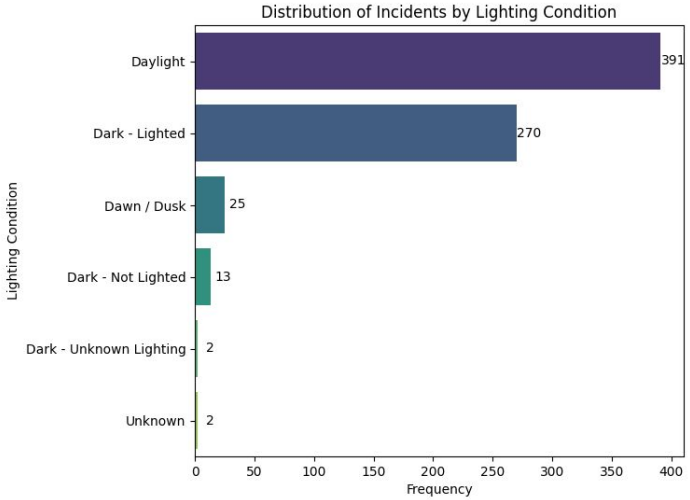


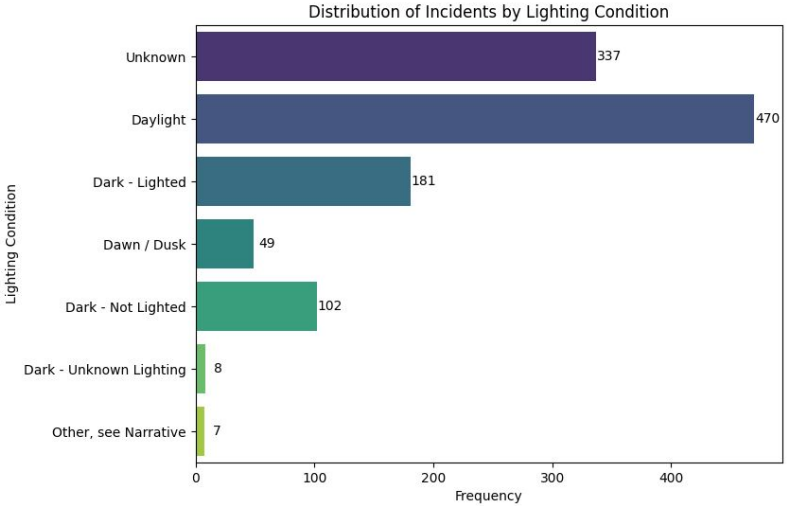
ADS, ADAS Level 2 Posted Speed Limit



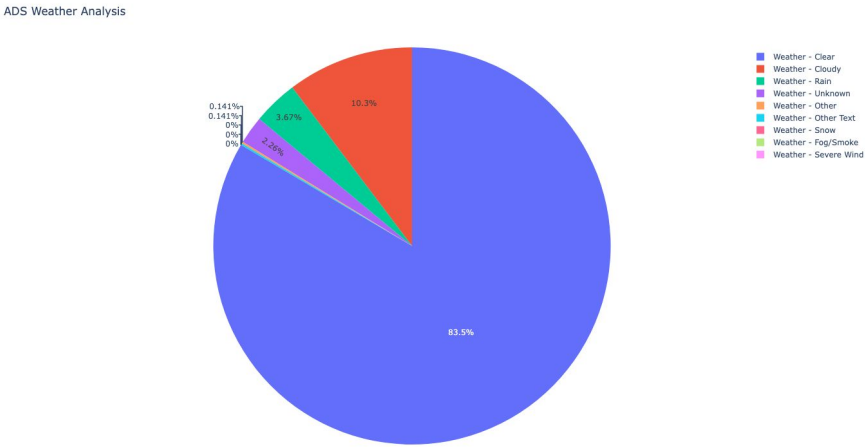


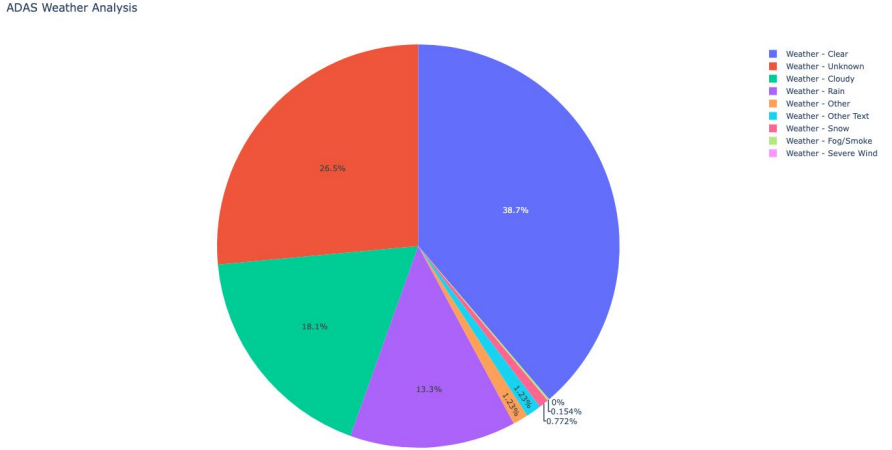
ADS, ADAS Level 2 Lightning





ADS, ADAS Level 2 Weather Analysis

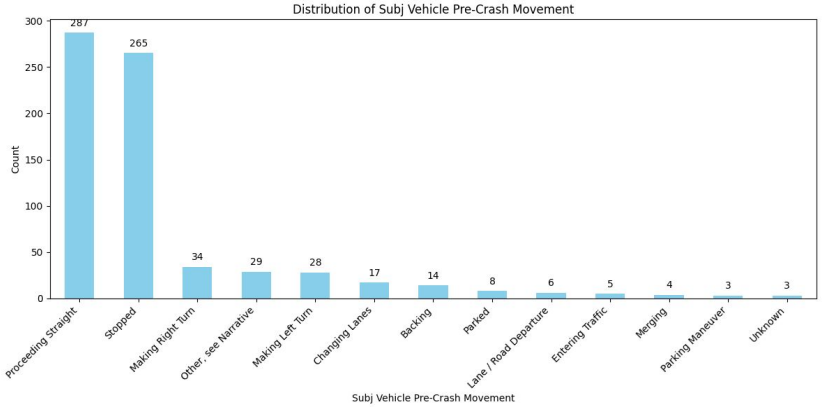


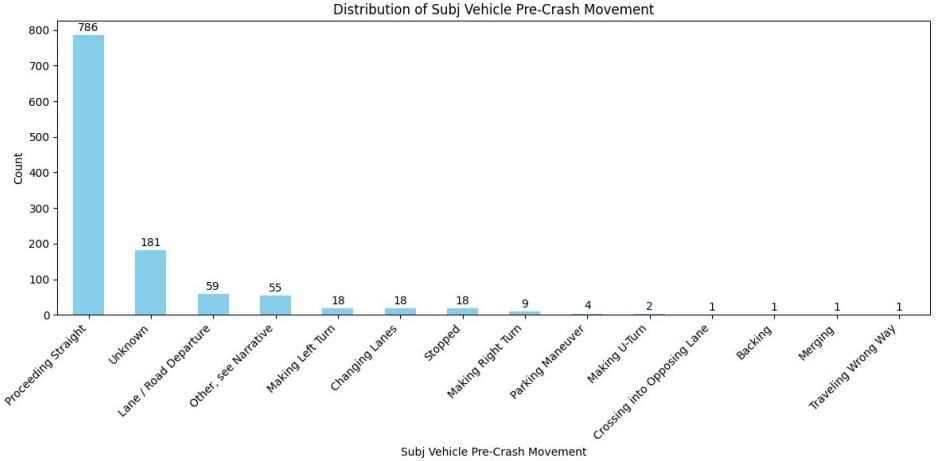


Data Visualization

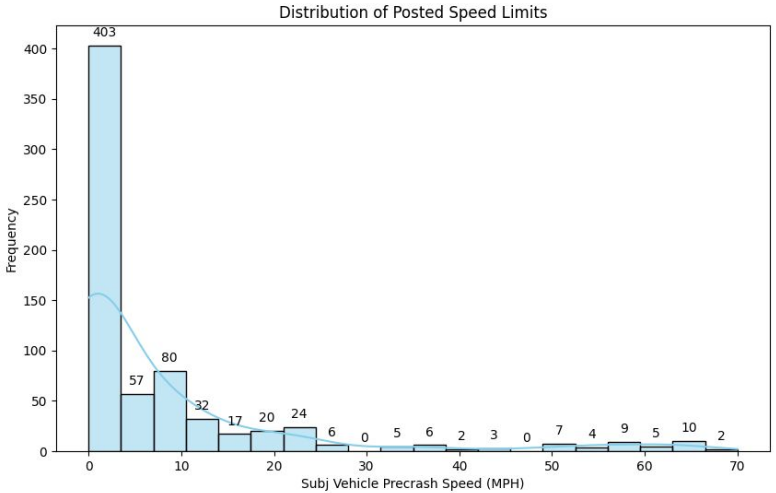
- Which?: reporting companies count
- Who?: Operator analysis
- Where?: city ,state
- When?: Incident time
- What?: weather, road type, mileage, speed limit
- **Why?: precrash movements**

ADS, ADAS Level 2 Subject Vehicle pre-crash movement

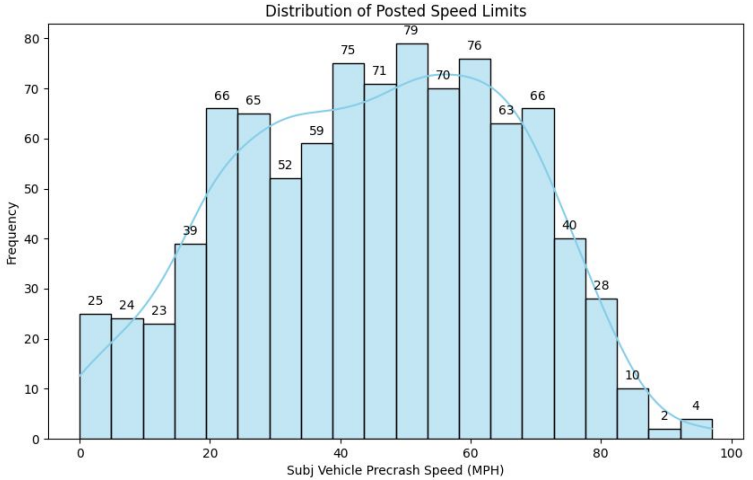




ADS, ADAS Level 2 Subj Vehicle Pre-Crash Speed Analysis

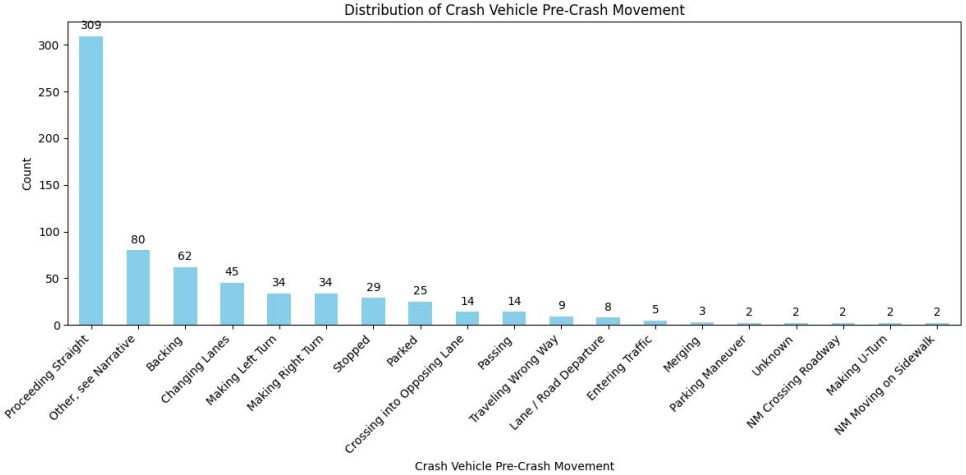


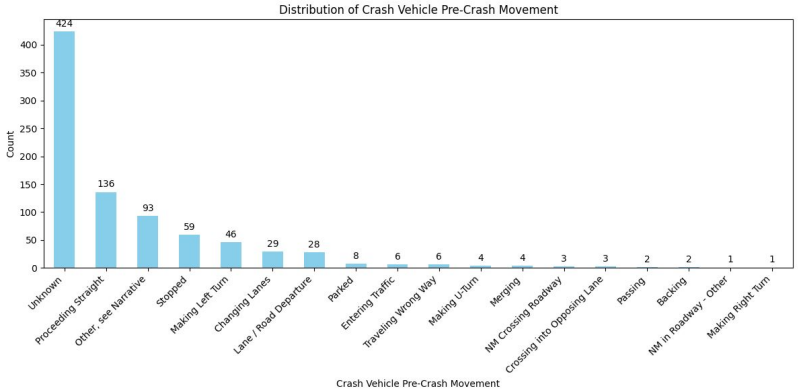

```
Mean Speed Limit: 8.091040462427745
Median Speed Limit: 1.0
Standard Deviation of Speed Limit: 14.482849003872918
```




```
Mean Speed Limit: 45.18143009605123
Median Speed Limit: 46.0
Standard Deviation of Speed Limit: 20.92710069931179
```

ADS, ADAS Level 2 Crash Partner pre-crash movement

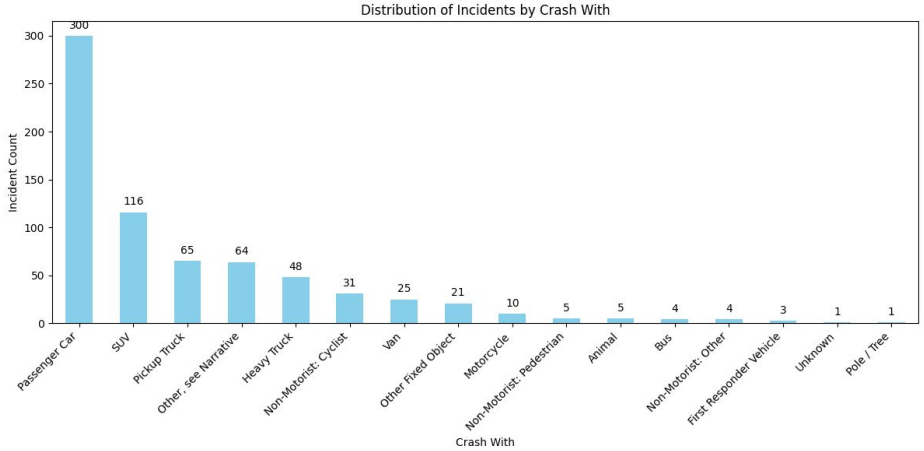


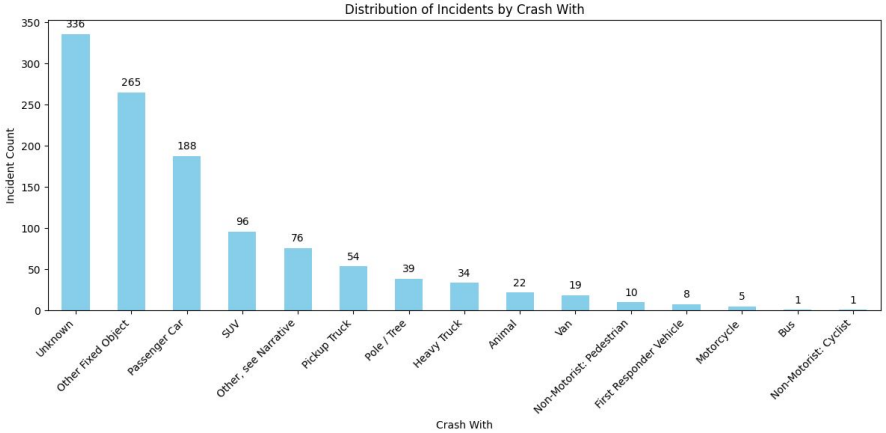


Data Visualization

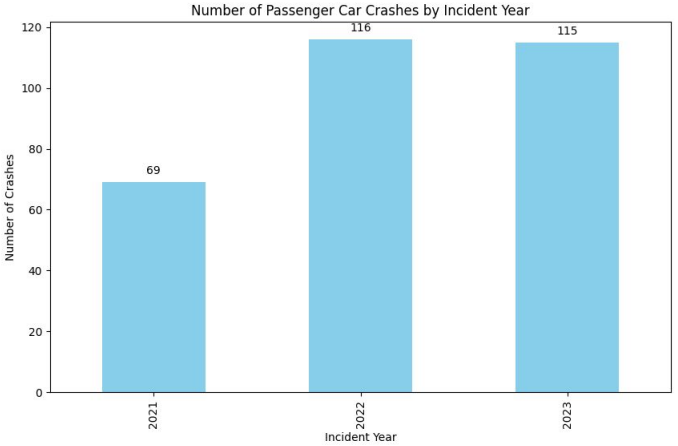
- Crash Analysis
  - Crash with?
  - Injury Analysis
  - Contact Area
  - Safety Analysis

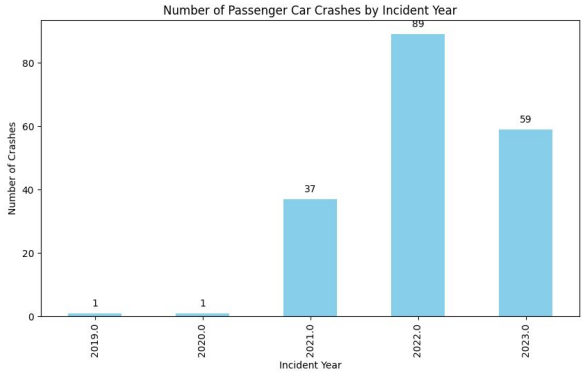
ADS, ADAS Level 2 Crash With?



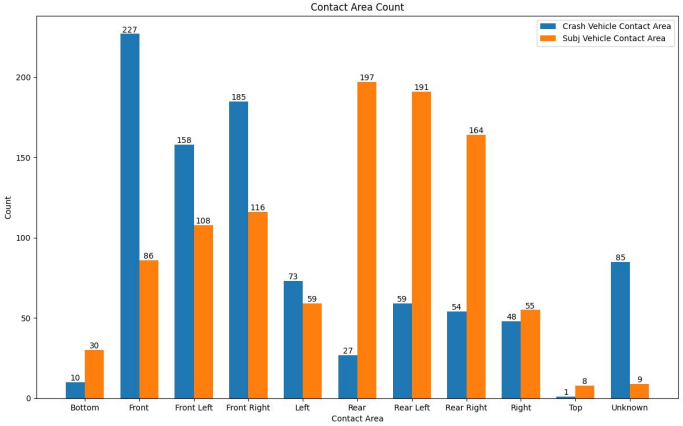


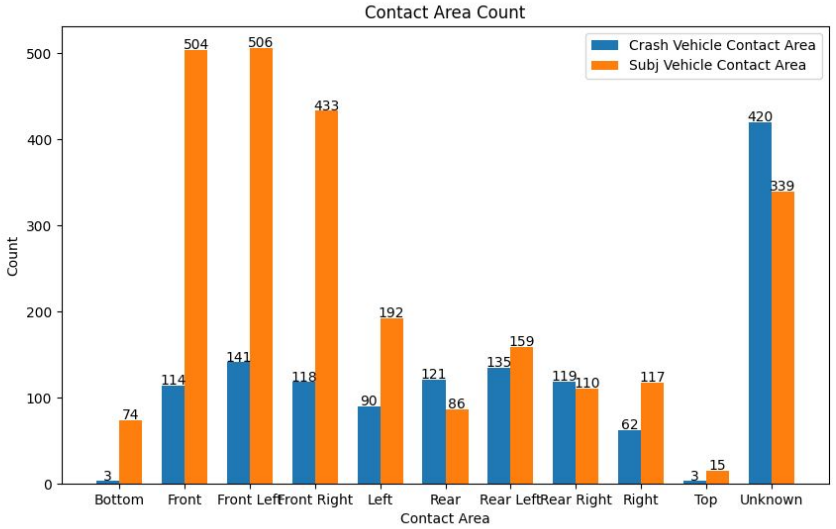
ADS, ADAS Level 2 Crash with Passenger Vehicle



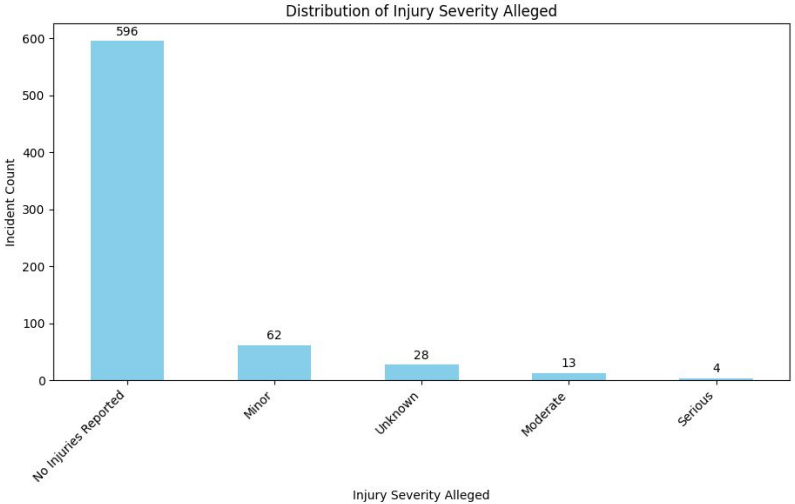


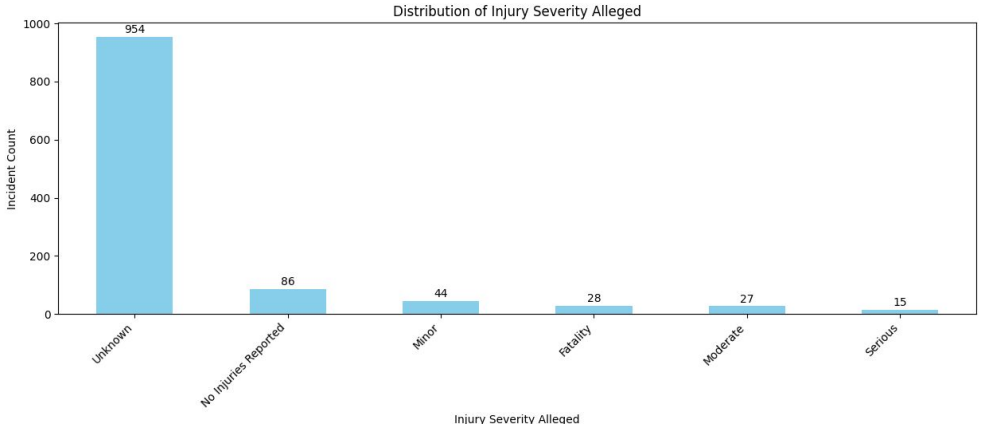
ADS, ADAS Level 2 Contact Area



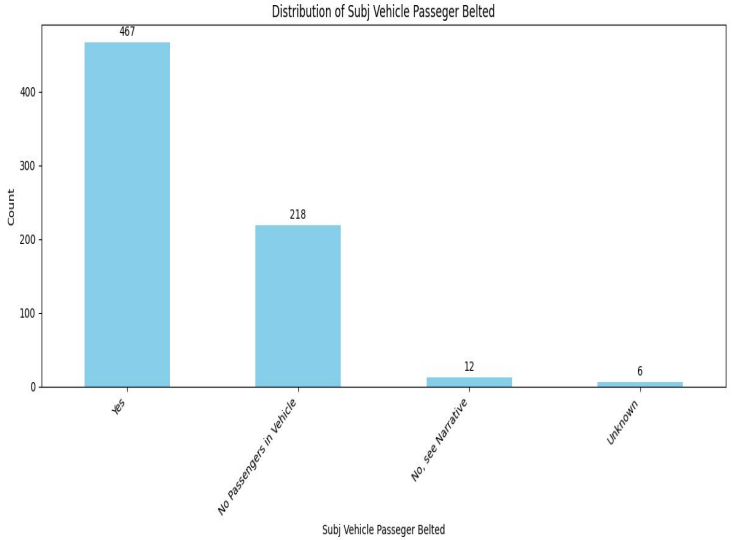


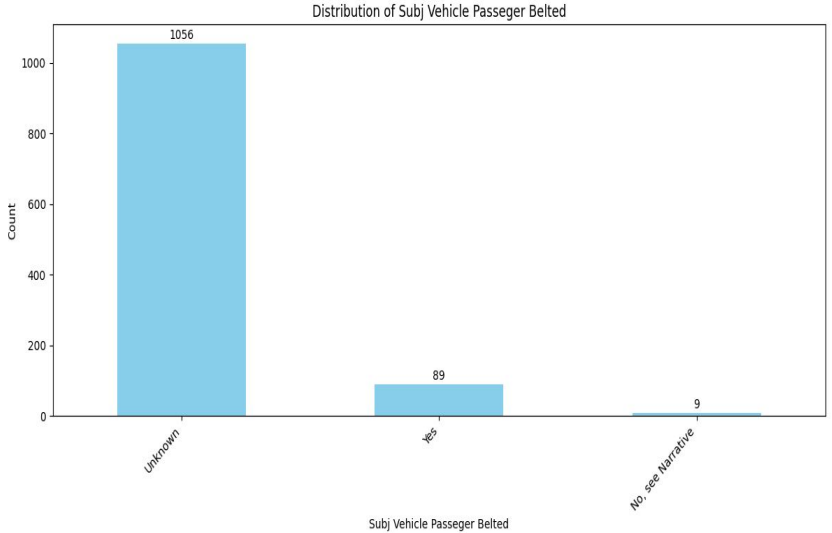
ADS, ADAS Level 2 Injury Analysis





ADS, ADAS Level 2 Subject Vehicle Safety





ADS, ADAS Level 2 Facts

1. SV Precrash Speed is higher than Posted Speed Limit
   a. Number of accidents: 237 in ADAS
   b. Number of accidents: 4 in ADS
2. SV Precrash Speed (MPH) and Posted Speed Limit (MPH) has stronger the correlation.
3. There are version control of the submitted reports highest goes still 5
4. ADS has 57 accidents due to sensor issues.
5. There different report types that the reporting companies has to update.

ADS, ADAS Level 2 and Human Driven vehicle Analysis in AZ (21-23)

- Count of accidents
  - ADS:96
  - ADAS:25
  - Human Driven: 200k
- Peak Location
  - ADS: Phoenix, Tempe
  - ADAS:Phoenix
  - Human Driven: Flagstaff
- Highest Road Type
  - ADS: Dry surface street
  - ADAS: Unknown surface Highway / Freeway
  - Human Driven: Dry surface level road

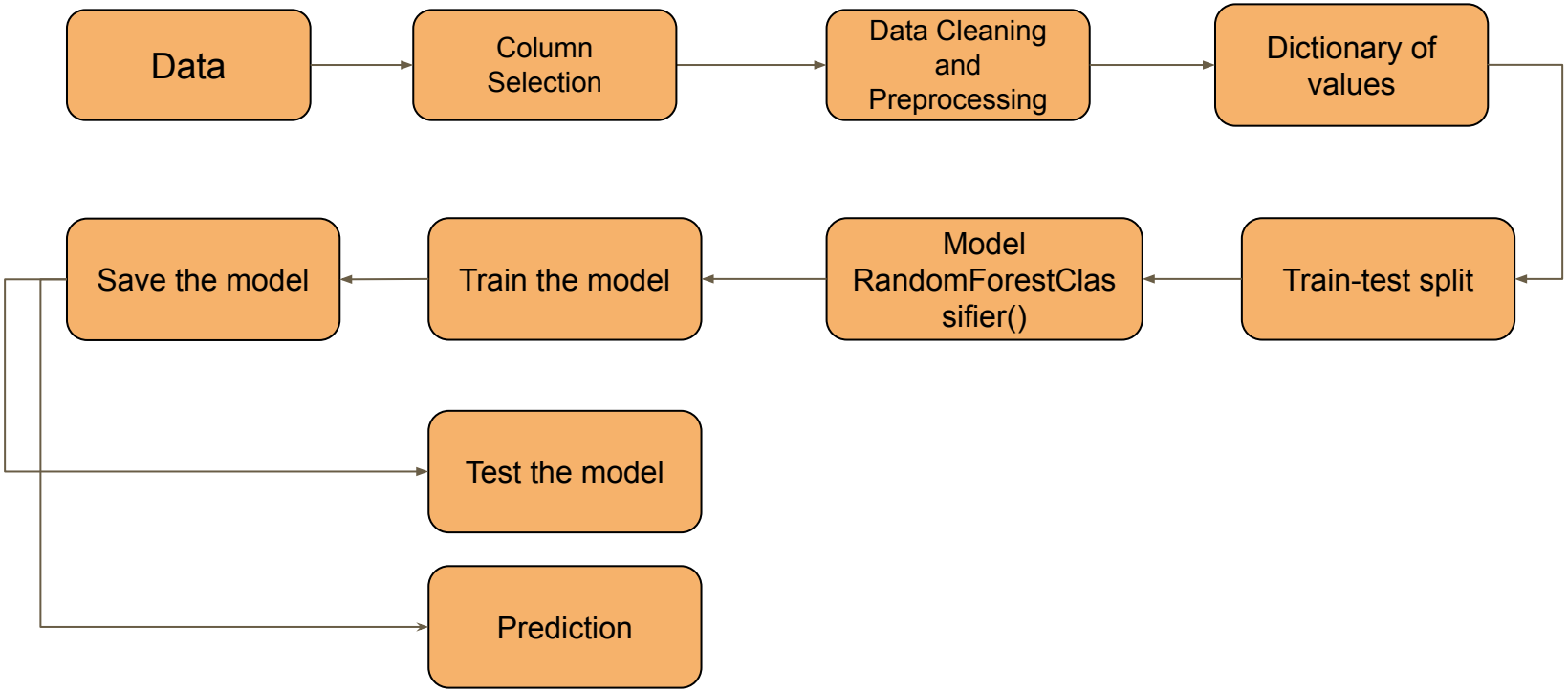
ADS, ADAS Level 2 and Human Driven vehicle Analysis in AZ (21-23)

- No of speeding Accidents
  - ADS: 0
  - ADAS: 2
  - Human Driven:1538
- Crash With?
  - ADS:Passenger Car
  - ADAS:Unknown
  - Human Driven:Passenger Car
- Peak time of Accidents
  - ADS: 1PM
  - ADAS: 12 AM
  - Human Driven: 3-4 PM

ADS, ADAS Level 2 and Human Driven vehicle Analysis in AZ (21-23)

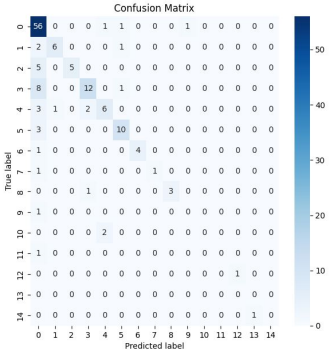
- Weather
  - ADS: clear
  - ADAS: clear
  - Human Driven: clear
- Lightning
  - ADS:Daylight
  - ADAS:Daylight and Unknown
  - Human Driven: Daylight

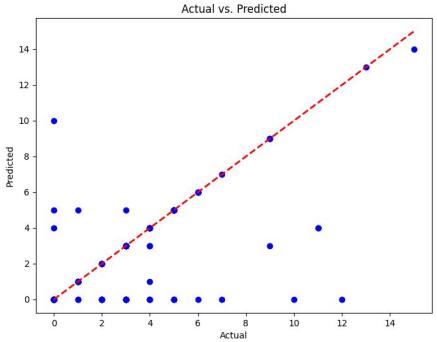
Prediction of Crashing Partner/ Vehicle: Flow Chart



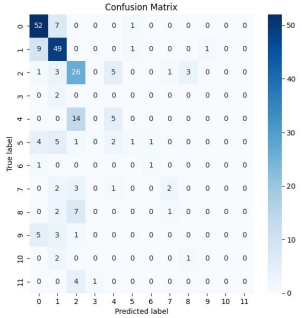
Prediction of Crashing Partner/ Vehicle: Accuracy

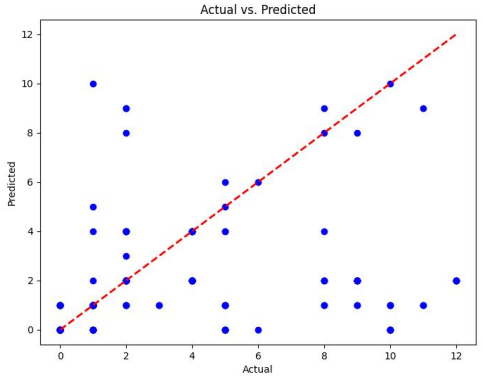
ADS Pred Accuracy: 74%



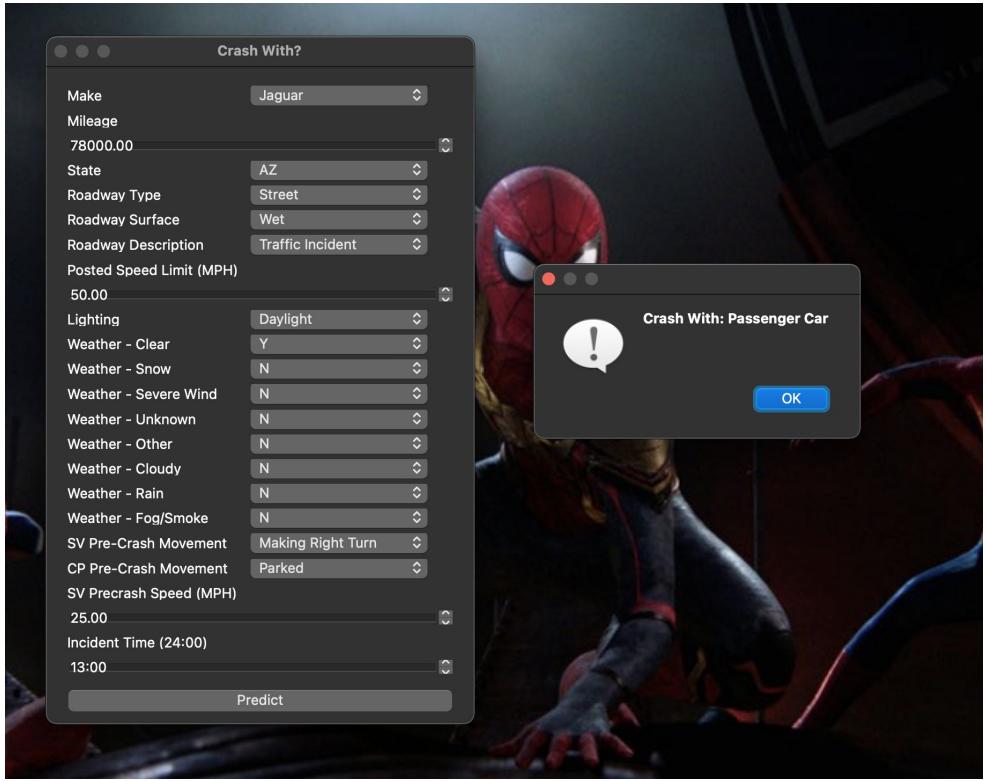


ADAS Pred Accuracy: 60%





Prediction of Crashing Partner/ Vehicle: GUI



Future work

- GUI user friendly
- Prepare model for different use case
- Collection of more data
- Trying using different ML algorithms for increasing the accuracy
- Exploring complex queries and analysis.
- Add more fields to the to fetch more precise data of incident.

Takeaways

Data analytics

ML models

Exploring complex queries

References

1. https://azdot.gov/mvd/services/statistics/arizona-motor-vehicle-crash-facts
2. https://www.nhtsa.gov/laws-regulations/standing-general-order-crash-reporting#overview
3. https://link.springer.com/article/10.1023/A:1010933404324
4. https://www.riverbankcomputing.com/software/pyqt/

Vote Of Thanks

Prof. Brendan Russo, NAU

Prof. Jeffrey Wishart, ASU


Q&A