Chicago Car Crashes:

Analyzing the Causes of The City's Traffic Accidents

Sameeha Ramadhan

May 2021

INTRODUCTION

- Traffic safety is a shared responsibility, crashes are **preventable**.
- The city of Chicago implemented the **Vision Zero** commitment in an effort to curb car crash injuries and fatalities.
- Building predictive models can analyze the cause of crashes and assist the city in taking correct measures to keep them from occurring.

VISION ZERO CHICAGO

PRESENTATION OUTLINE

- The data source and how it was filtered
- Observations made through analysis
- The most effective model for predicting the causes of car crashes
- Recommendations based on results

OBTAINING & PROCESSING THE DATA

The Data Source

- The data was sourced from the city of Chicago's website.
- The datasets are three: traffic crashes - crashes, traffic crashes - vehicles, and traffic crashes - people

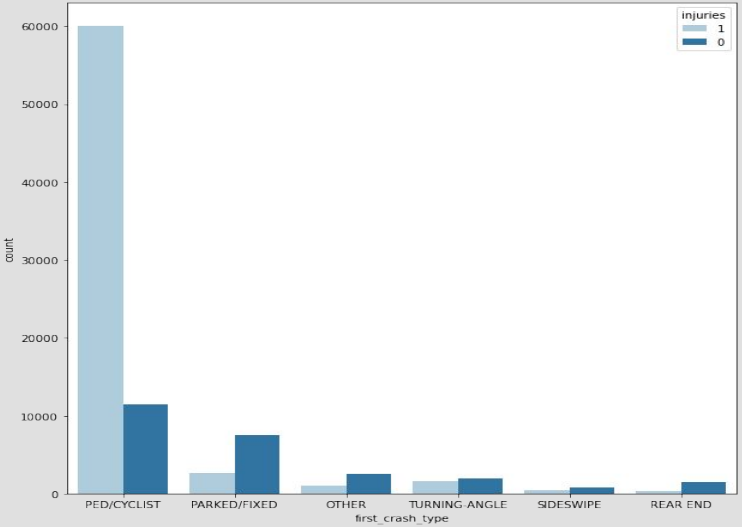
cityofchicago.org

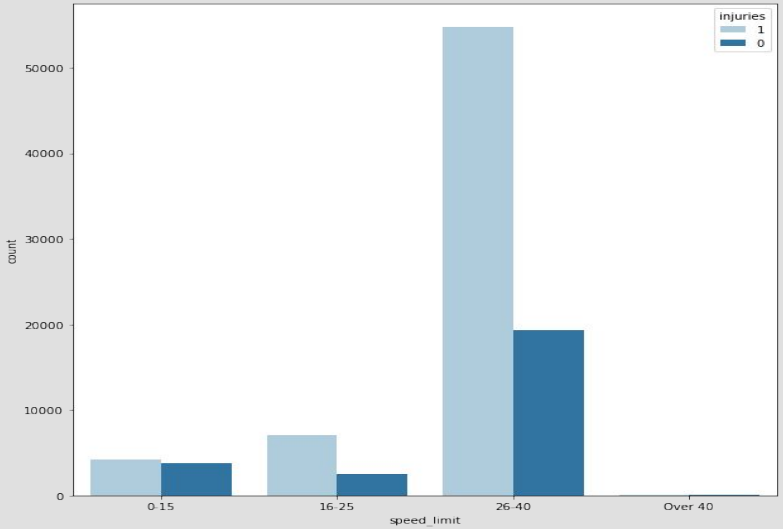
Cleaning the Data

- The datasets were filtered to 2021 only.
- 'Injuries' was the target
- This resulted in over 92,000 crash incidents for the model

EXPLORING THE DATA

- Most car accident related injuries are a result of drivers colliding with pedestrians and cyclists.

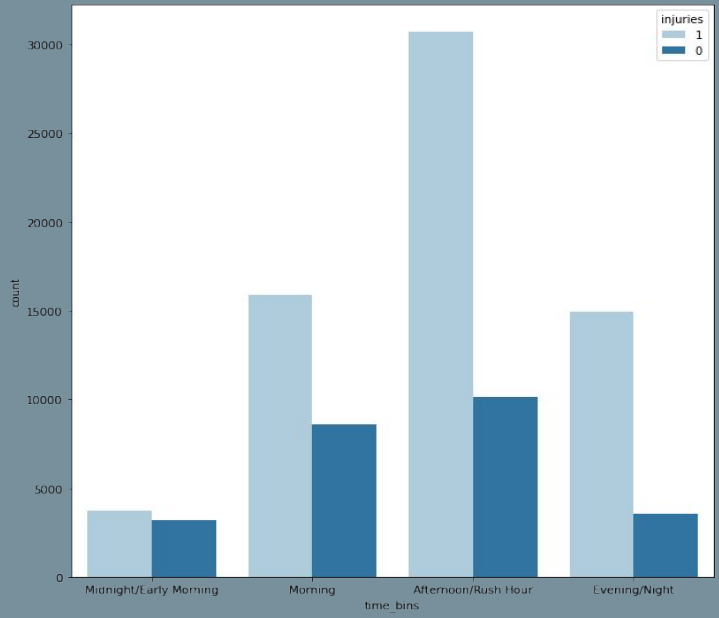


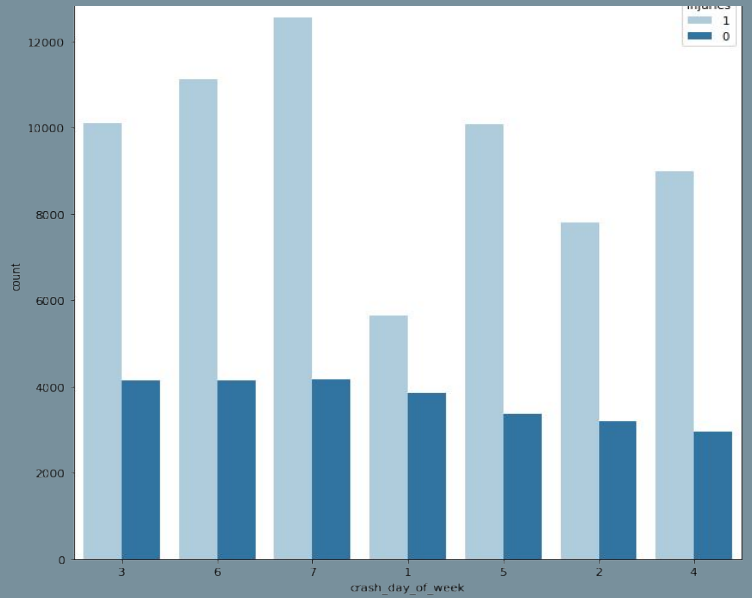


- The majority of car crashes and their resulting injuries occur in areas whose speed limits are between 26 and 40 mph.

EXPLORING THE DATA

- Most car accidents occur during afternoon rush hour.





- The majority of car crashes and their resulting injuries occur on Saturday, followed by Friday.

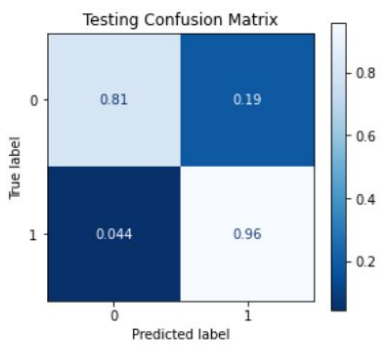
MODELING THE DATA

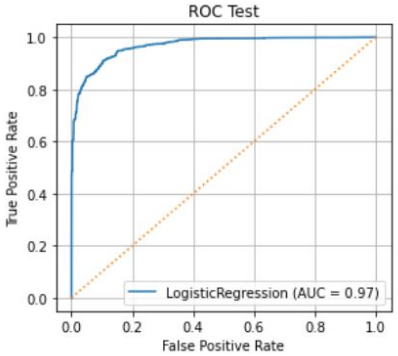
- The target variable was ‘injuries_total’ which was binned and renamed to ‘injuries’.
- Multiple models were tested, including KNN and Decision Trees.
- The Logistic Regression model had an accuracy of 92%.

```
------------------------------------------------------------
LOGISTIC REGRESSION Testing Classification Report
------------------------------------------------------------
              precision    recall  f1-score   support

           0       0.88      0.81      0.84      6413
           1       0.93      0.96      0.94     16608

    accuracy                           0.92     23021
   macro avg       0.90      0.88      0.89     23021
weighted avg       0.91      0.92      0.91     23021
```





CONCLUSIONS

Due to the model's accuracy rate, I am confident in the following:

- Most injuries result from collisions between **drivers and pedestrians or cyclists**.
- Accidents and injuries occur most often **in the presence of traffic signals**.
- The majority take place in the **afternoon or during rush hour** as well as on Saturdays.
- Most occur in speed limit zones between **30-40 mph.**

RECOMMENDATIONS

Non-Motor Protection	• Install cyclist friendly lanes • Designate more pedestrian walking areas
Rush Hour	• Lower speed limits • More control in accident prone speed zones
Traffic Flow	• Expand two-way roads • Install median/dividers
Requirements	• To wear bright/reflective clothing when dark • Classes on traffic safety

THANK YOU

Email: shramadhan@gmail.com
GitHub: @samtuleen
LinkedIn: linkedin.com/in/sameeha-ramadhan