PREDICT THE PRIMARY CONTRIBUTORY CAUSE OF A CHICAGO CAR ACCIDENT

BY OMARA WALDEA

PROBLEM

◈ Over 2000 people, including drivers, passengers, pedestrians, and cyclists, will be involved in over 1000 car accidents over a four-day period. Of them, 45% will have minor to fatal injuries. Predicting the causes of accidents can help prevent injuries.

BUSINESS VALUE

◈ Safer streets for drivers, passengers, pedestrians, and cyclists

◈ Less work for emergency responders and officers

◈ Reduced traffic and accidents

◈ Shorter travel times

◈ Improved safety for Chicago residents

◈ Satisfied drivers.

DATA

DATA SOURCE: CHICAGO WEBSITE

- Vehicle Information

- People Information

- Traffic Crashes Information

METHODOLOGY

1. Combine Data:

- Merge the Vehicle Data, Driver/Passenger Data, and Crash Data based on a common identifier (e.g., accident ID).

2. Feature Engineering:

- Create features such as weather conditions, time of day, road type, vehicle type, and driver age group.

- Transform categorical variables into numerical values if necessary.

3. Handling Missing Values:

- Address missing values by imputation or by removing incomplete records, depending on their impact on the dataset.

METHODOLOGY

 4. Target Variable:

 - Define the target variable as the primary contributory cause of the accident.

 - Consider consolidating categories with few samples to ensure the model has enough data to learn from.

5. Data Splitting:

- Split the data into training and testing sets.

MODEL

◈ MODEL SELECTION: DECISION TREES

RECOMMENDATION

- Pedestrian/Cyclist wear contrasting colors

- Pedestrian/Cyclist are more aware during afternoon/rush hour

- The city can increase traffic signs/signals and lanes that are more ped/cyclist friendly

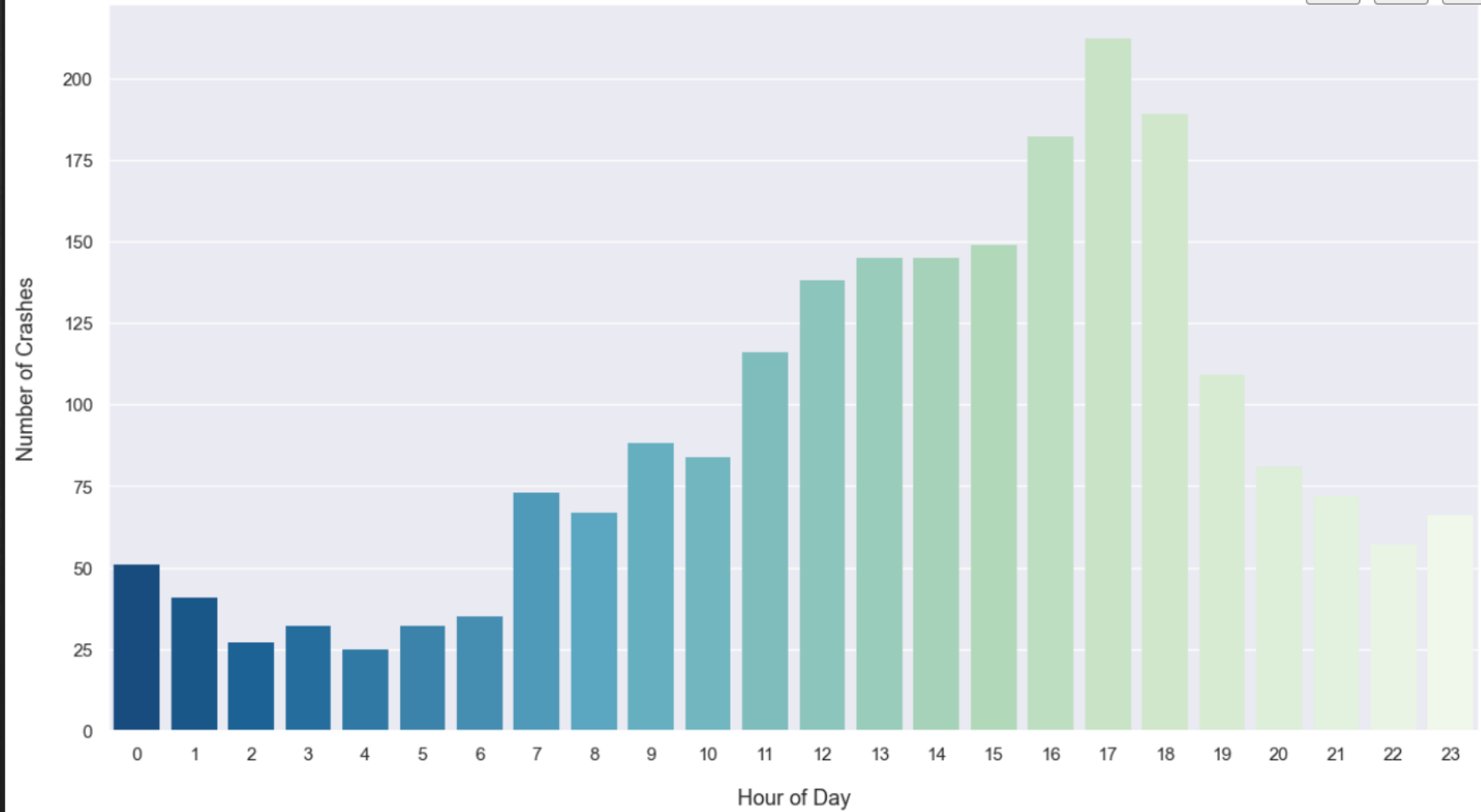- The city lowers the speed limit during afternoon/rush hour

Primary Contributing Cause of Reported Crashes in Chicago (2018 - 2020)
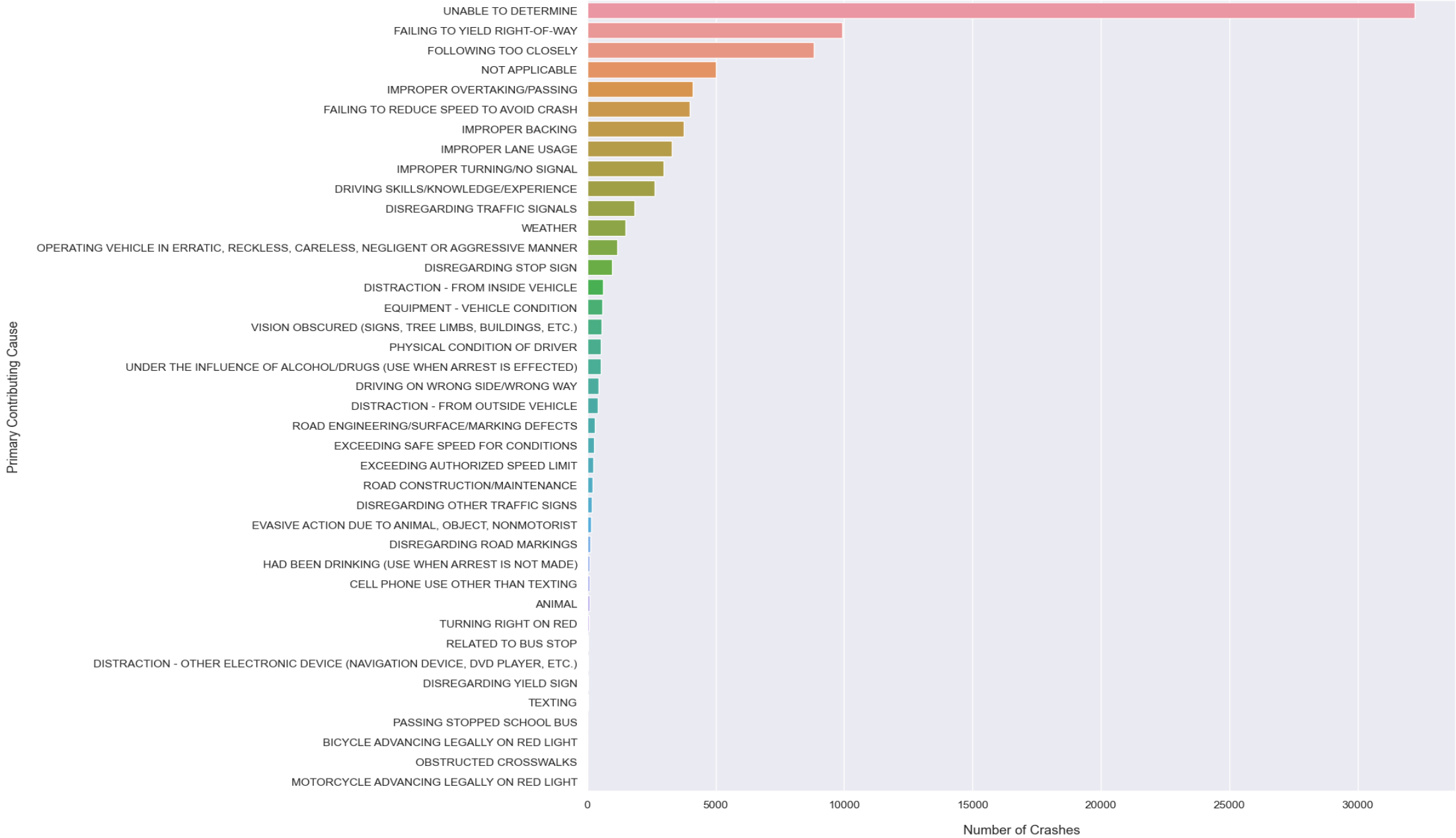
Bar Plot of INJURIES_TOTAL vs WEATHER_CONDITION
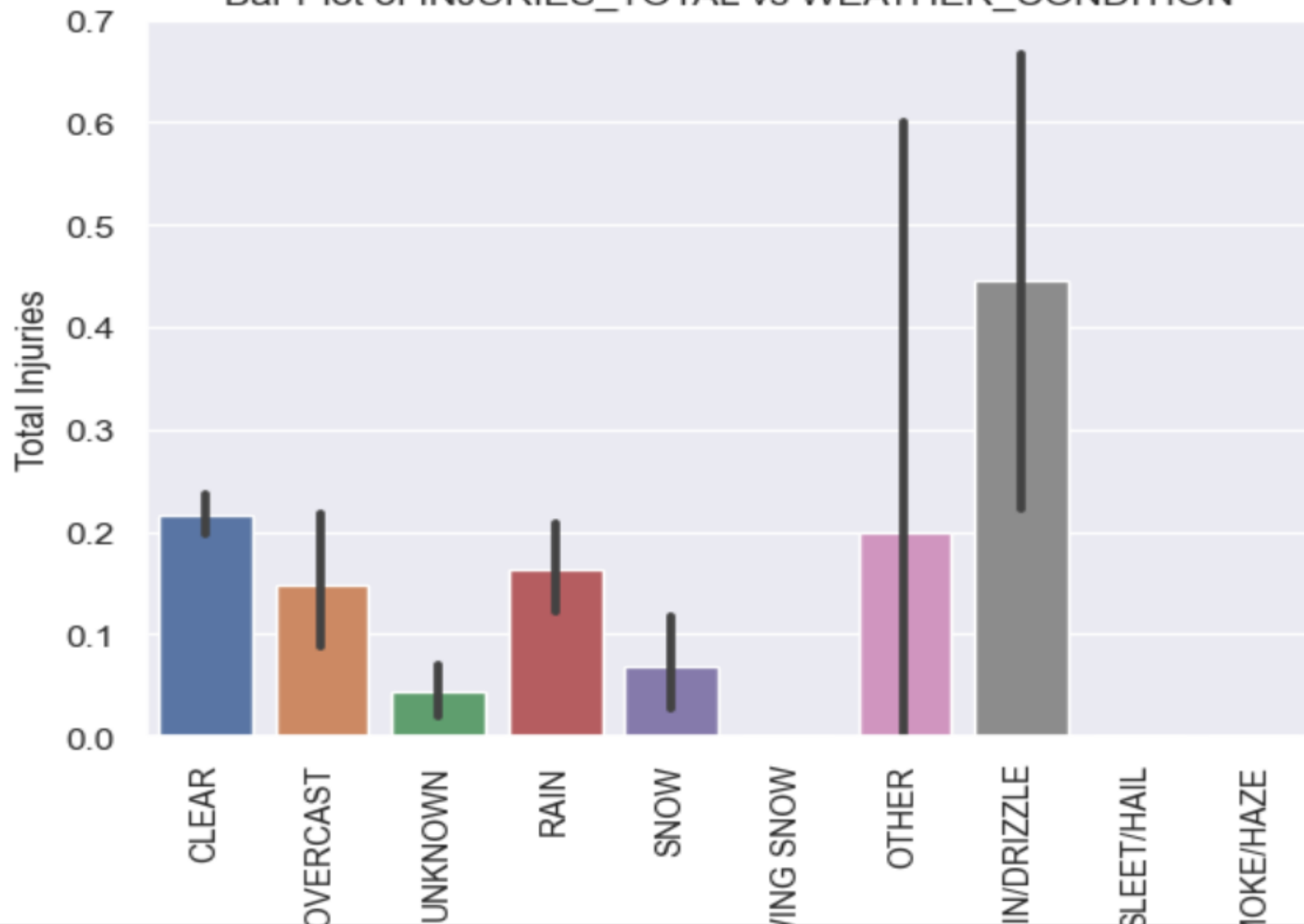
Hourly Number of Reported Crashes in Chicago (2018 - 2020)

THANK YOU ☺