

The background of the slide is a photograph of the front of a car, showing the hood, headlights, and grille. The image is heavily overlaid with a semi-transparent purple color, which serves as a backdrop for the white text.

COURSERA CAPSTONE

# Predicting Car Crash Severity under Non-Ideal Weather and Road Conditions

---

August 2020 | Marc Castro



# Can we predict death?

## LOOKING INTO WHAT DETERMINES DEATH

Every year, approximately **6 million car crashes** occur in the U.S. where :

- **72%** of these crashes result in **property damage**
- **27 %** resulting in injuries ranging from **bruises to life-changing disabilities,**
- and **6%** resulting in **DEATH**

Factors such as road traction, rain, lighting, and even the type of traffic in the area may play a role in predicting these accidents and could potentially serve as a tool to avoid them.



# Data Acquisition & Cleaning

## **PARSING** | CLEANING | CREATING

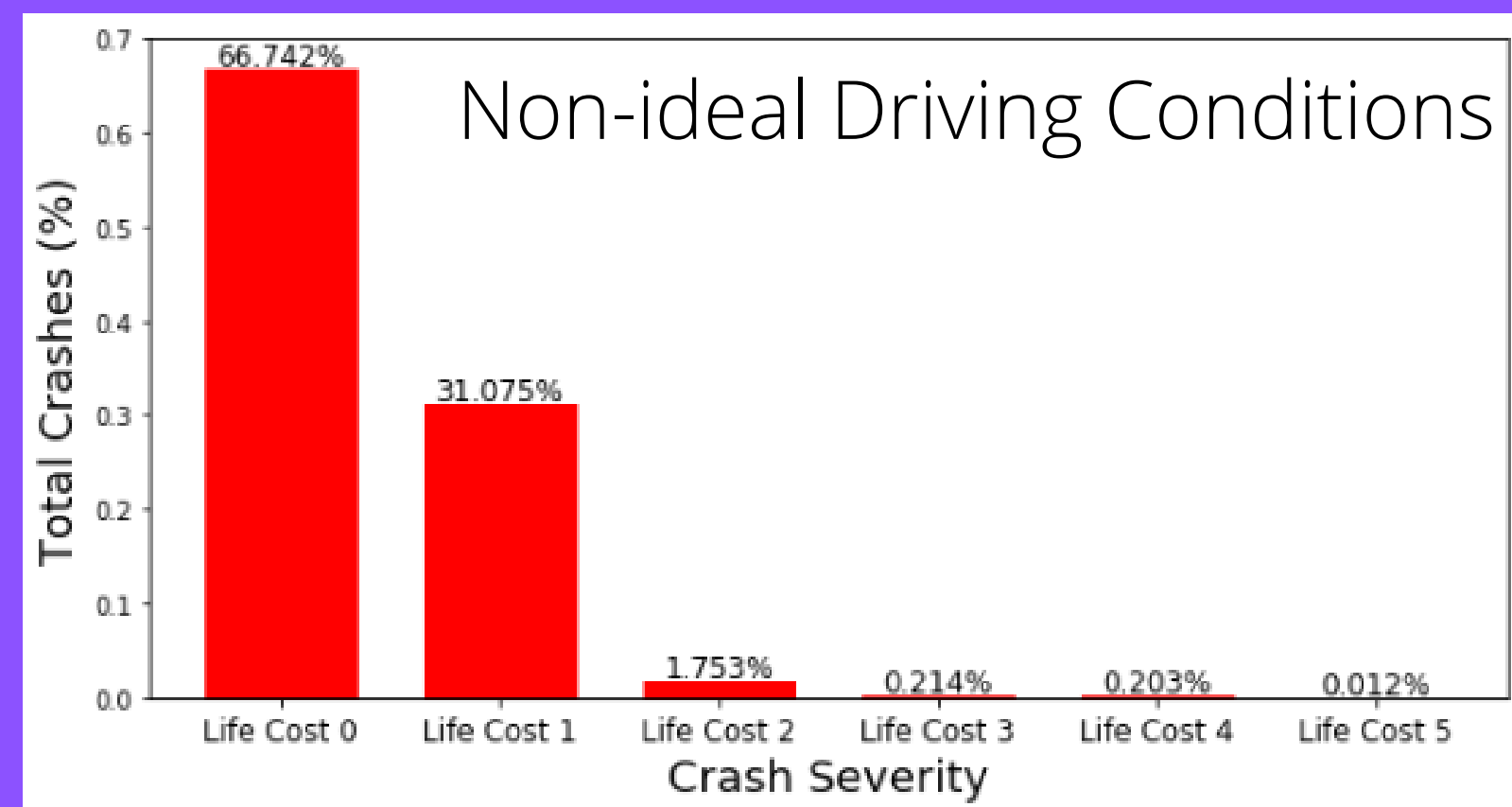
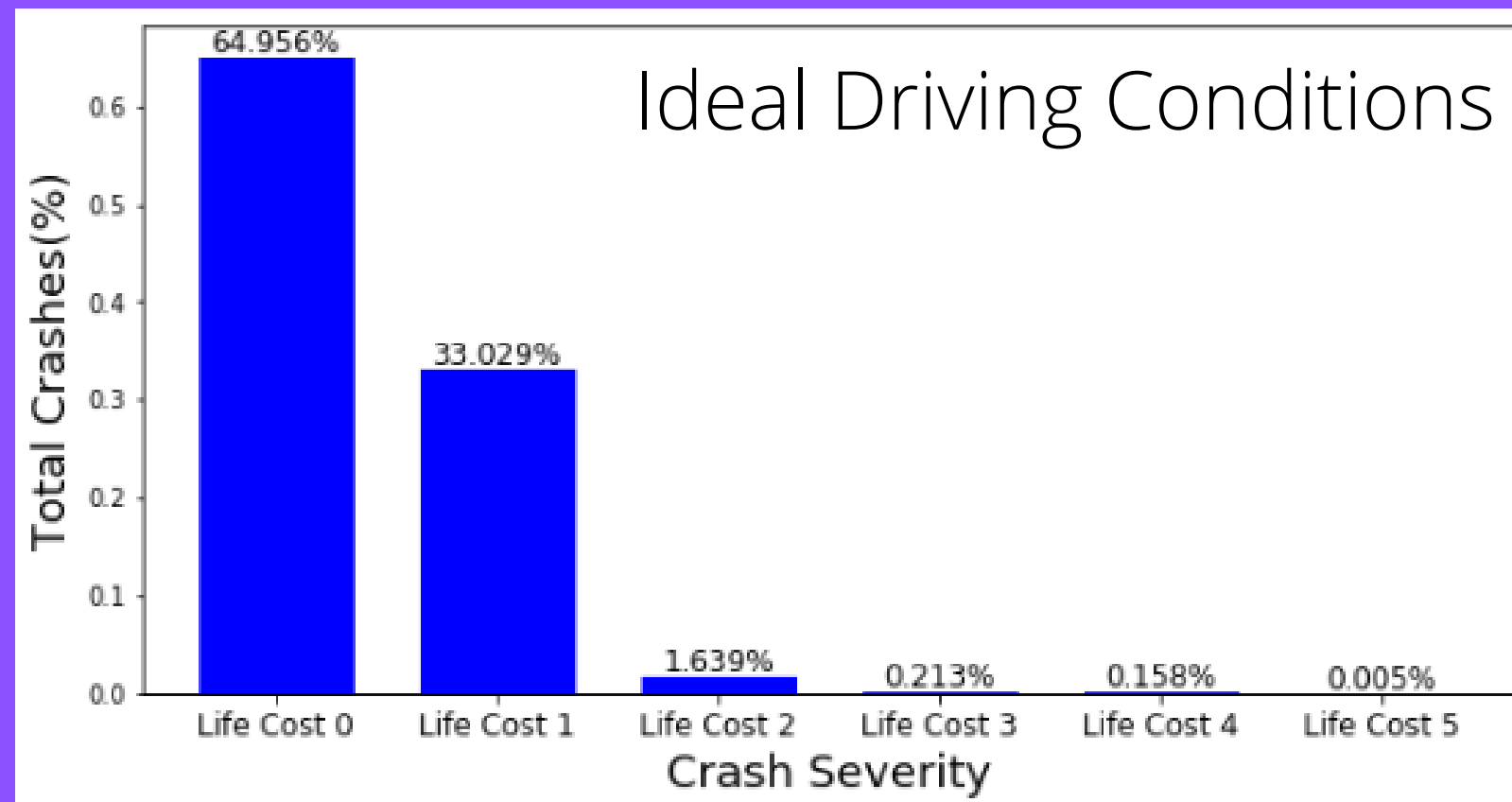
- Car Crash data for Seattle USA from 2004 to Present was retrieved from <https://data.seattle.gov/Land-Base/Collisions/9kas-rb8d>
- Raw Dataset contained 220,937 rows and 40 features.
- A total of 7 features were retained from the raw dataset subject for data cleaning.

## **PARSING** | **CLEANING** | CREATING

- Null values and Vague Values were removed by removing their corresponding row entirely.

## **PARSING** | CLEANING | **CREATING**

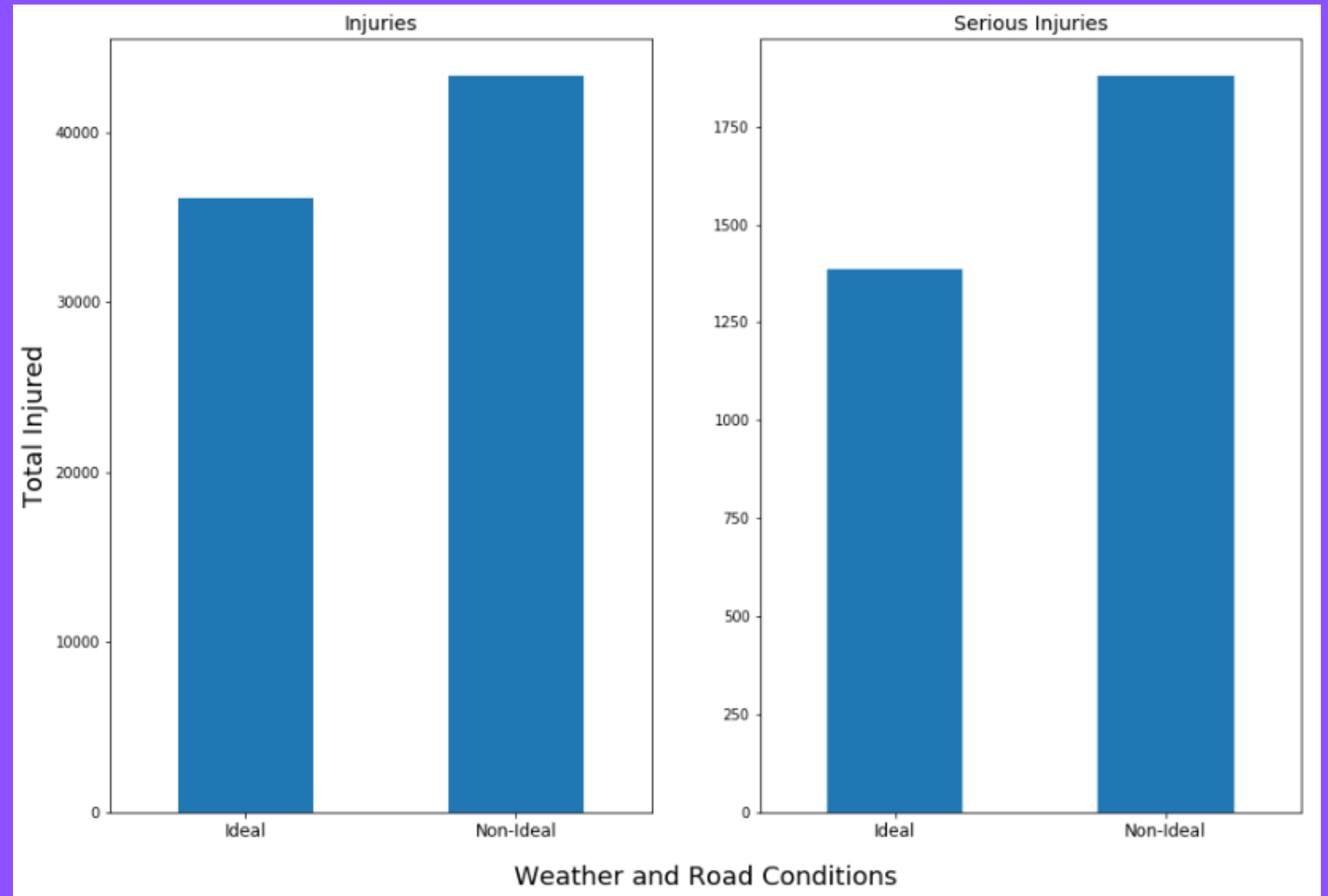
- Categorical values were binned with respect to similar highly similar types
- Target variable LIFECOST was generated by using a scoring system that scaled nominal values involving property damage, injury, and death



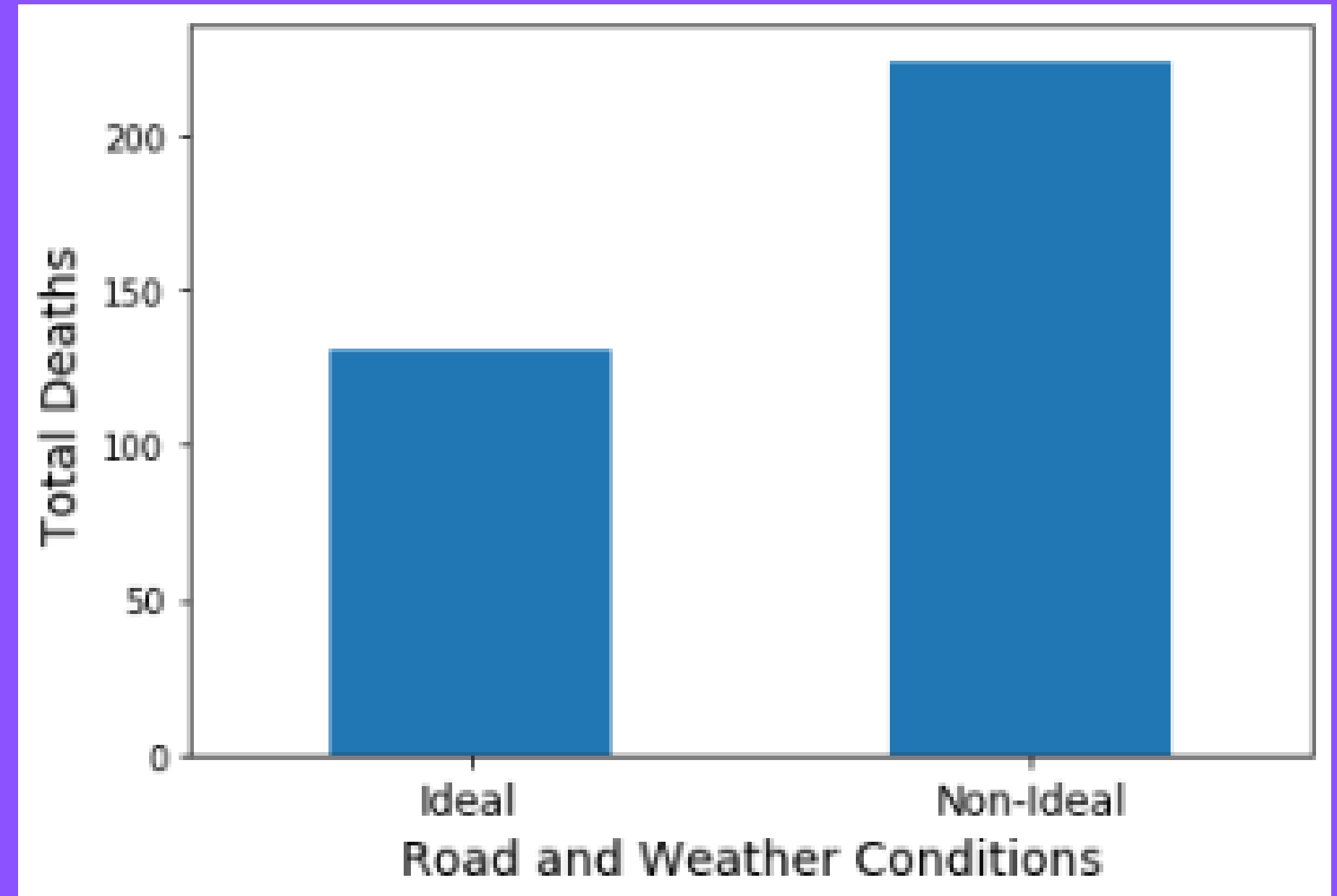
**EXPLORATORY DATA ANALYSIS**

**MORE SEVERE  
CRASHES OCCUR  
UNDER NON-IDEAL  
DRIVING CONDITIONS**

**THERE ARE MORE  
INJURIES AND  
DEATHS THAT  
HAPPEN IN  
CRASHES UNDER  
NON-IDEAL  
DRIVING  
CONDITIONS**



**THERE ARE MORE  
INJURIES AND  
DEATHS THAT  
HAPPEN IN  
CRASHES UNDER  
NON-IDEAL  
DRIVING  
CONDITIONS**

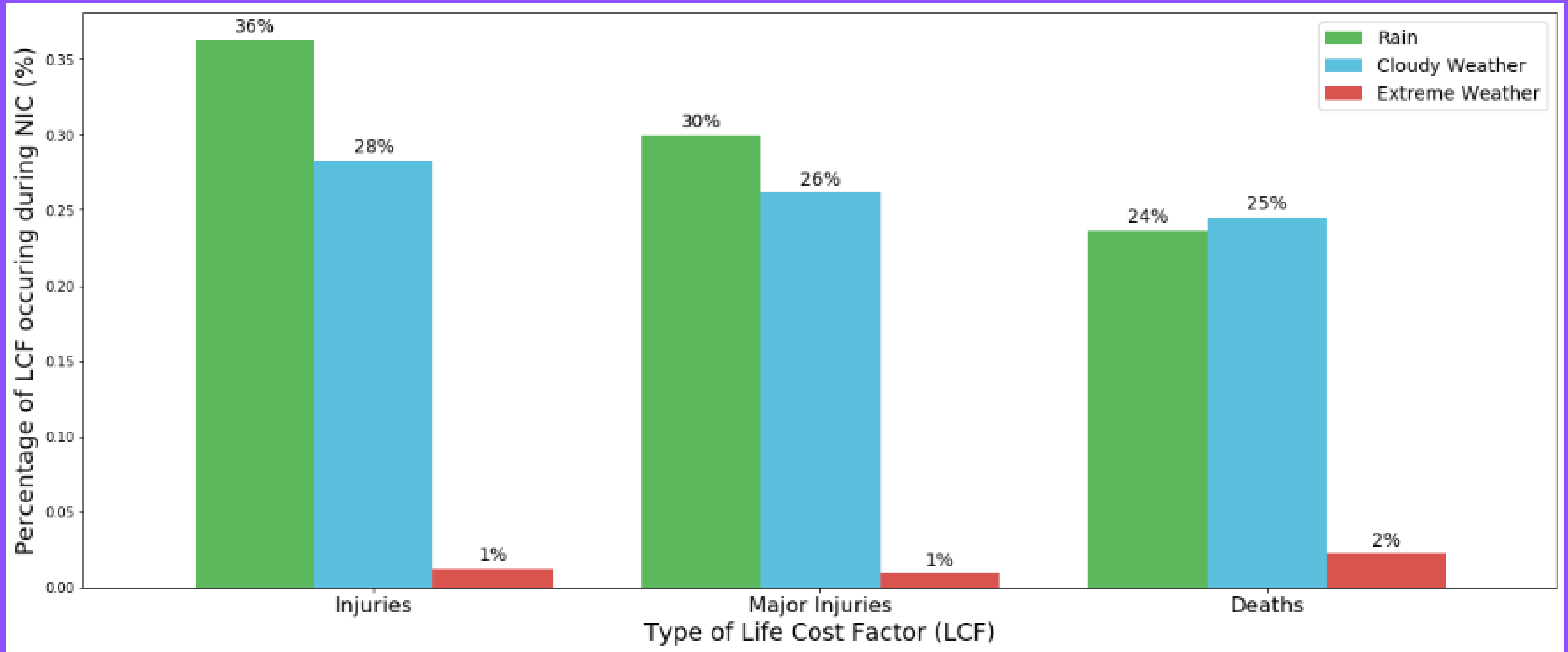


**DEATHS INCREASED**

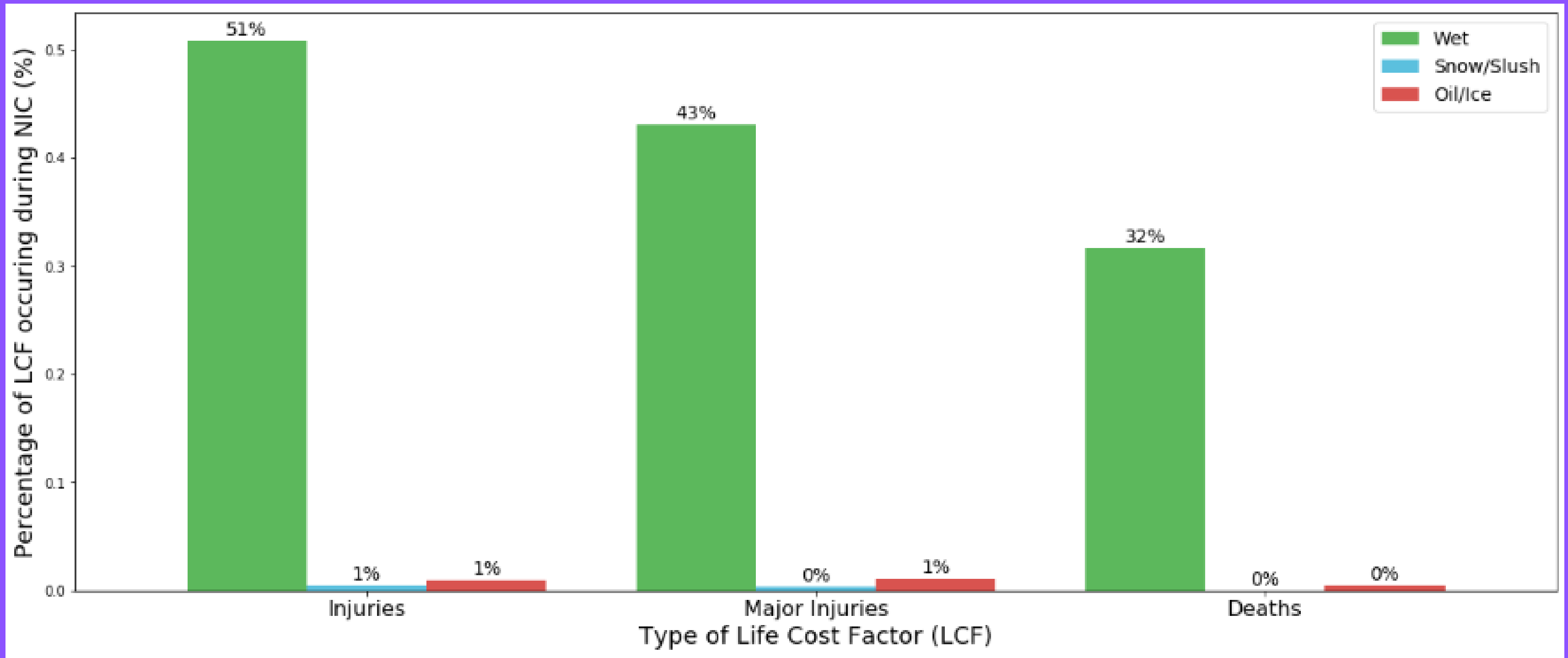
*26%*

There were generally more deaths involved in crashes during Non-Ideal Conditions

# MORE INJURIES & DEATHS OCCURED UNDER NON-IDEAL DRIVING WEATHER

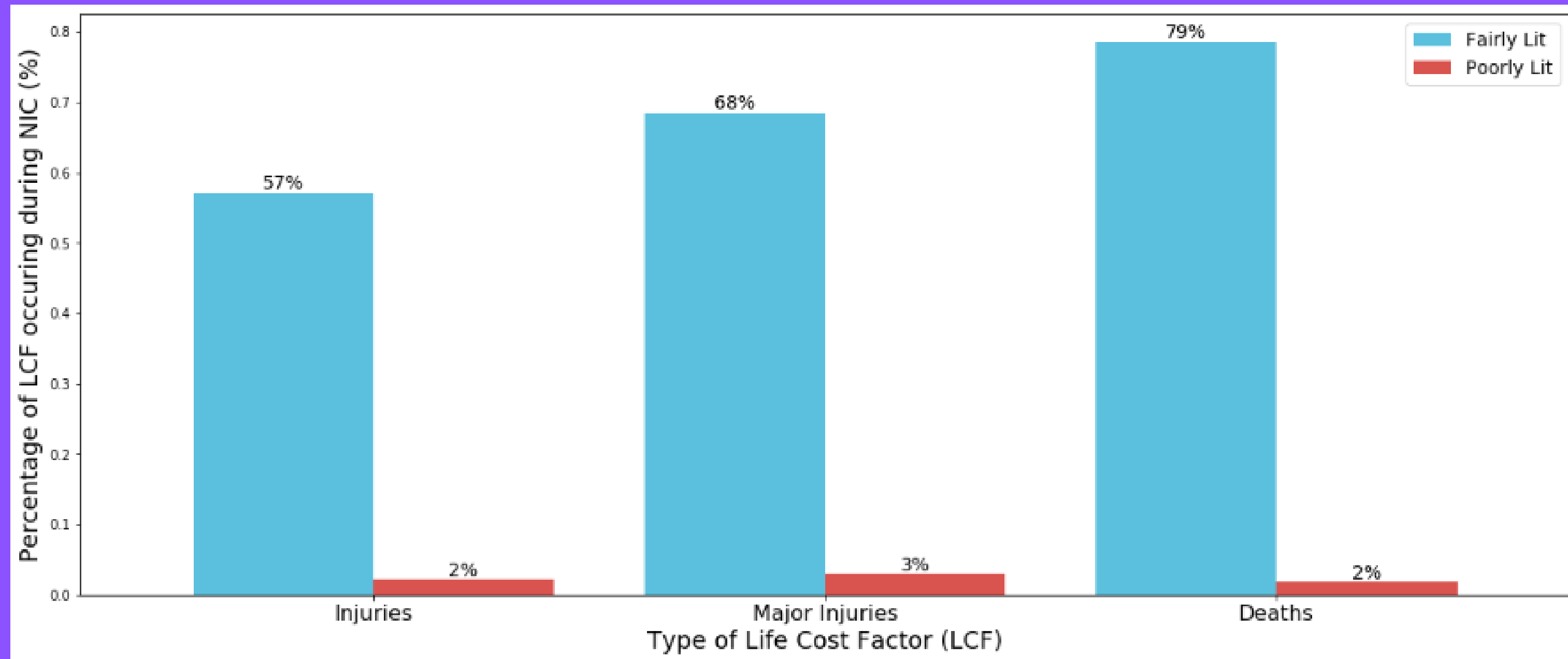


# LESS INJURIES & DEATHS OCCURED UNDER NON-IDEAL DRIVING ROAD CONDITIONS





# SURPRISINGLY, MORE INJURIES AND DEATHS OCCURED UNDER MODERATE LIGHT CONDITIONS COMPARED TO POOR LIGHT CONDITIONS



# CAR CRASH SEVERITY MODEL

## CLASSIFICATION ACCURACY

Classification of accidents based solely on categorical values of road condition, weather, and light condition performed poorly as class imbalance resulted to a large bias towards lower LIFECOST values which are significantly more.

## WHY THIS IS OKAY

We are not that concerned with the actual classification of data points, rather we are interested in the prediction probabilities generated by our trained model.  
Note that Logistic Regression was used as the final model.

## PREDICTION PROBABILITIES

In contrast to car severity classification, we are able to see the probability of each type of accident occurring based on weather and road conditions.

## RESULTS?

Our findings showed that some road and weather conditions actually increased the risk of death and multiple major injuries from occurring. Conversely, there were also some findings that indicate that some conditions actually reduced the risk of these high impact accidents.

# PREDICTION PROBABILITY FINDINGS

## CLOUDY WEATHER

There are generally more crashes involved with major injuries and less with death.

## EXTREME WEATHER

It is 6.29% more likely to be involved in accidents involving multiple major injuries and death in contrast to ideal driving conditions. While nearly double the likelihood of involvement with accidents that have multiple deaths.

## RAIN

It is 2.77% more likely to have more major injuries and a higher chance of death in contrast to ideal driving conditions.

## WET ROAD

You are more exposed to accidents involving major injuries than to those involving minor injuries.

## FAIRLY LIT ROAD

There is ~ 80% increase of the possibility that you are involved in a crash with multiple minor injuries and are less likely to be involved in more severe crashes.



## BUILT A MODEL

I was able to build a model that would determine the probability of a crash occurring with a respective crash severity using road conditions and weather conditions.



## IDENTIFIED CONDITIONS WITH HIGH PROBABILITY FOR SEVERE ACCIDENTS

Some conditions resulted in higher probabilities of major injuries and deaths such as extreme weather, rain, and wet roads.



## IDENTIFIED CONDITIONS WITH HIGH PROBABILITY FOR MINOR ACCIDENTS

Some instances that resulted in a lower probability for accidents involving multiple major injuries such as driving on fairly lit roads (Dusk, Dawn, Dark - with Street Lights On) in contrast to well lit roads( Mid-day light)



## FUTURE DIRECTIONS

find a dataset that contains specific nominal information on the weather and road conditions during the time of accident. Information such as wind speed, temperature, humidity, friction, and etc. could potentially determine key factors that result in these crashes.

# CONCLUSION AND FUTURE DIRECTIONS