Investigating Chicago Auto Incidents

By: Payson Chadrow

Goal

Identify leading factors of auto incidents to locate areas of improvement for safety

TABLE OF CONTENTS

01 DATA OVERVIEW

02 DATA EXPLORATION

03 CLASSIFICATION AND EFFECTIVENESS

04 RESULTS

05 THE ROAD AHEAD

DATA OVERVIEW

01

WHAT WE WORKED ON

OVERWHELMING AMOUNTS OF DATA

3 different datasets totaling at close to a million observations

INCORRECT DATA

Abundance of data that did not make sense in a realistic setting.

- over 1,000 lanes present

MEANINGLESS DATA

Features recording the same value for every entry

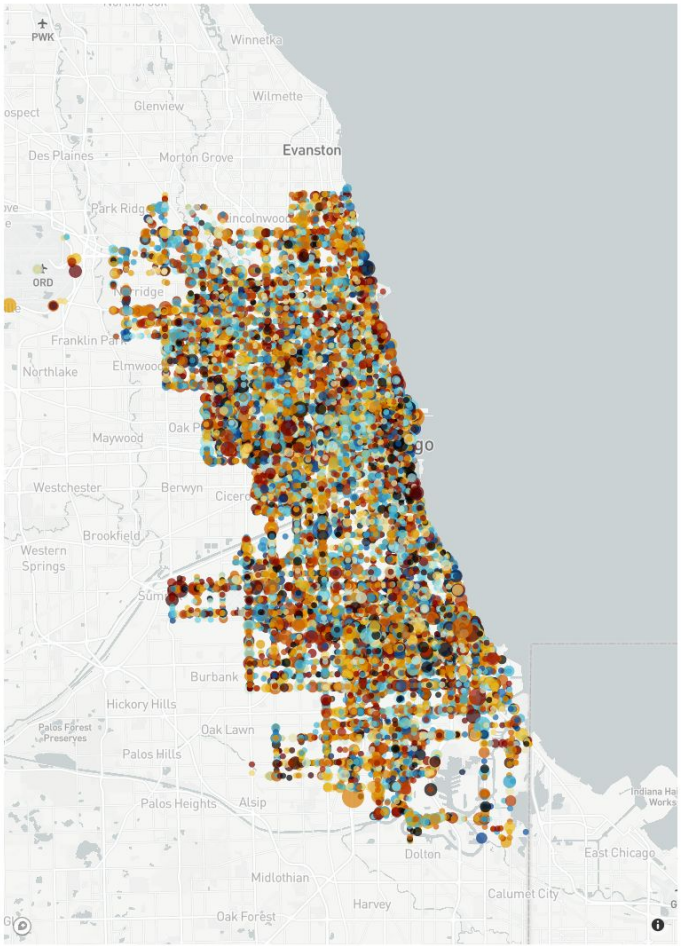
MISSING DATA

Crashes - >20%
People - >35%
Vehicles - >70%

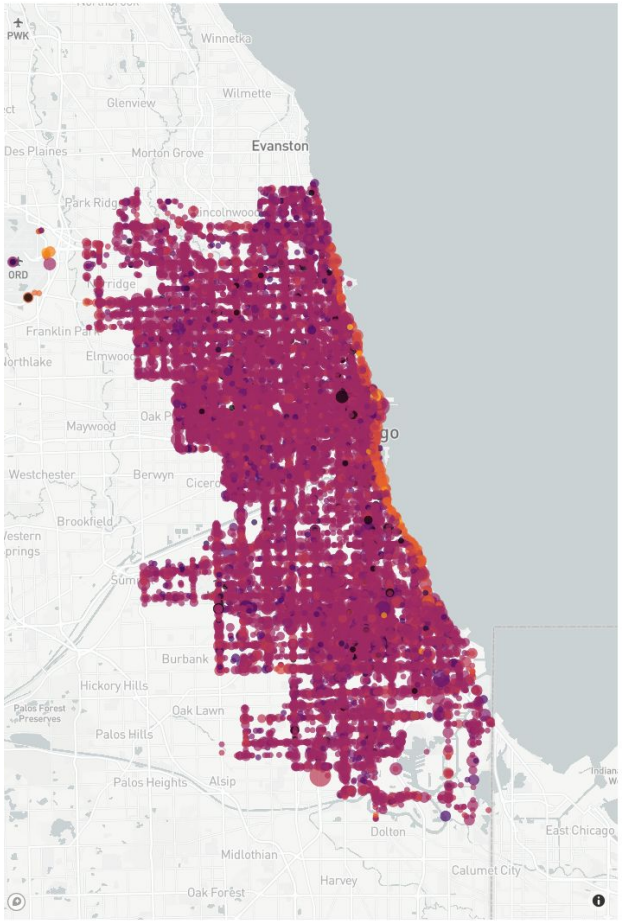
DATA EXPLORATION

02

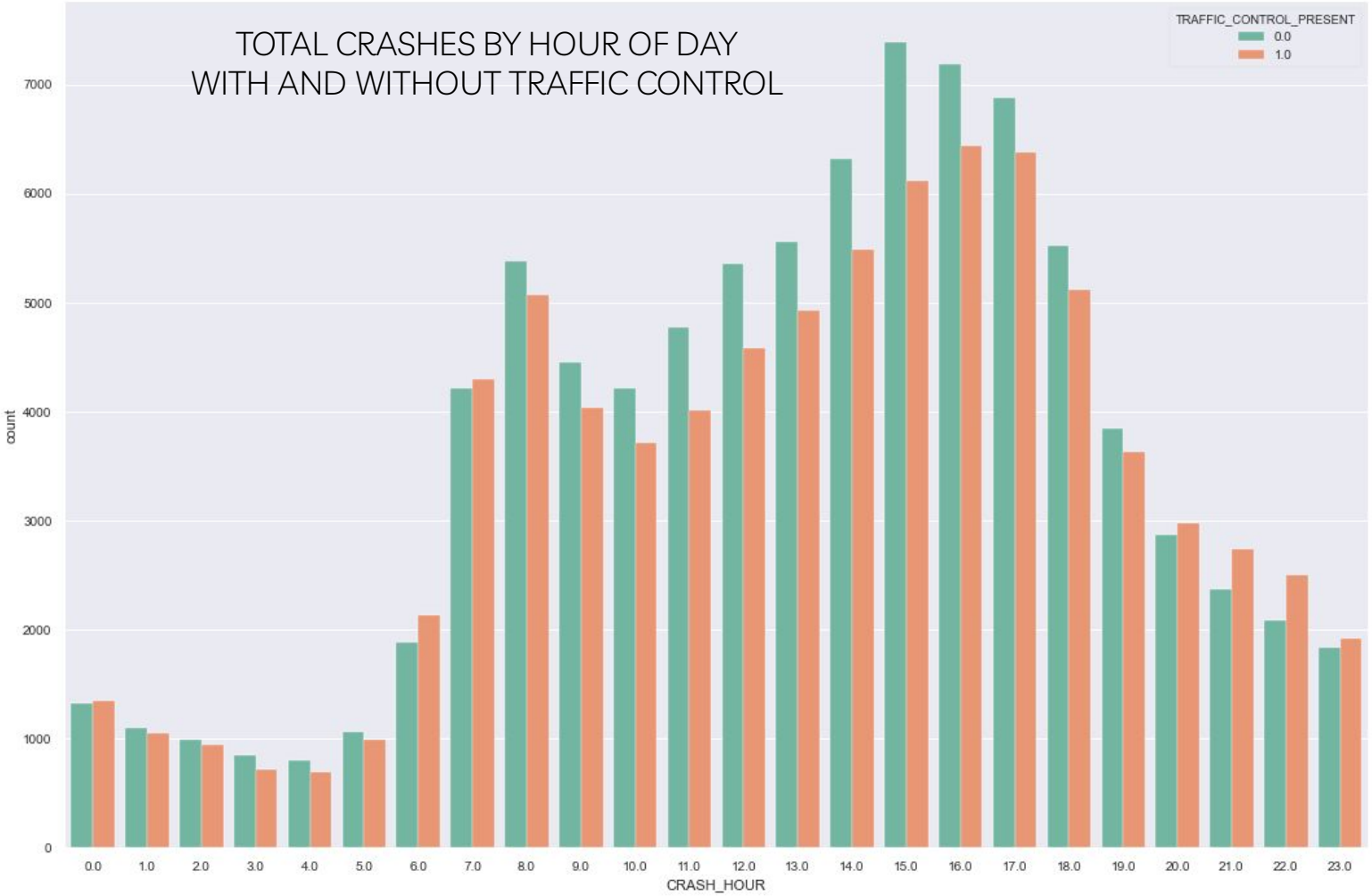
Chicago Crash Locations - Time of Day



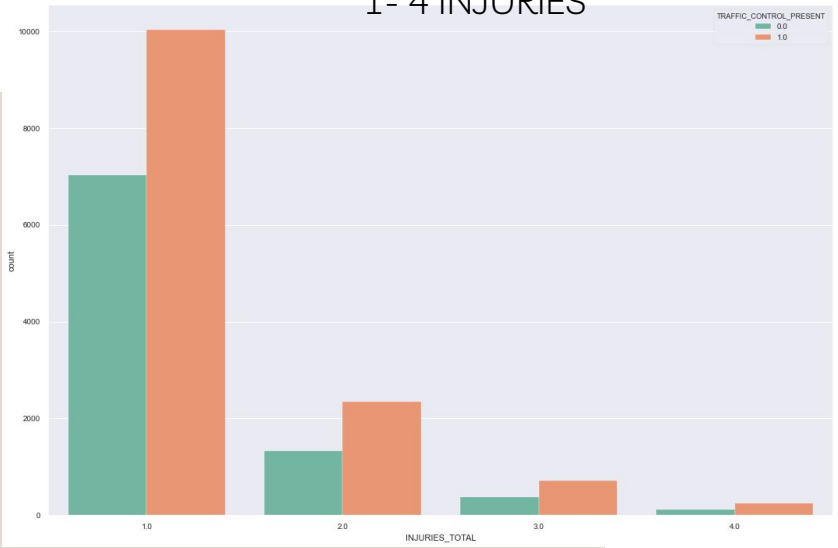
ago Crash Locations - Filtered



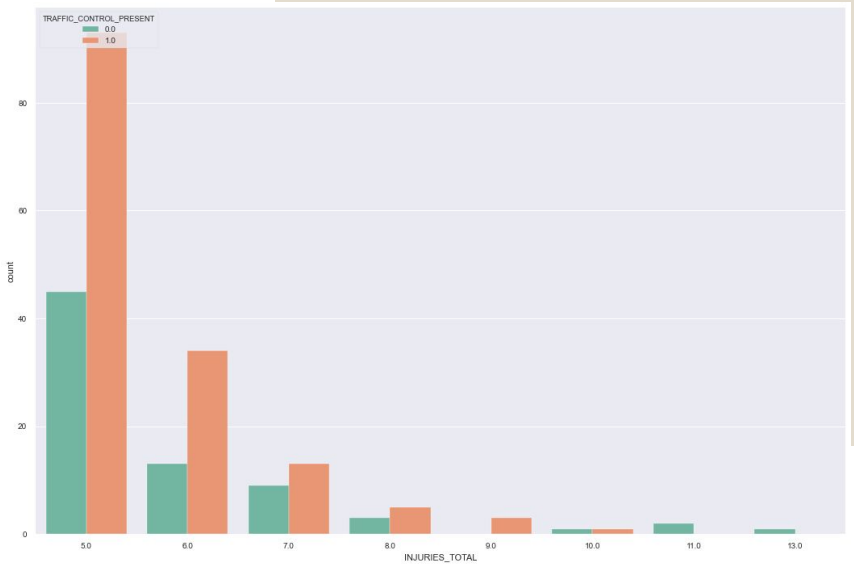
TOTAL CRASHES BY HOUR OF DAY WITH AND WITHOUT TRAFFIC CONTROL



INJURY DISTRIBUTION BY TRAFFIC CONTROL IN THOUSANDS
1- 4 INJURIES



INJURY DISTRIBUTION BY TRAFFIC CONTROL IN TENS
5+ INJURIES



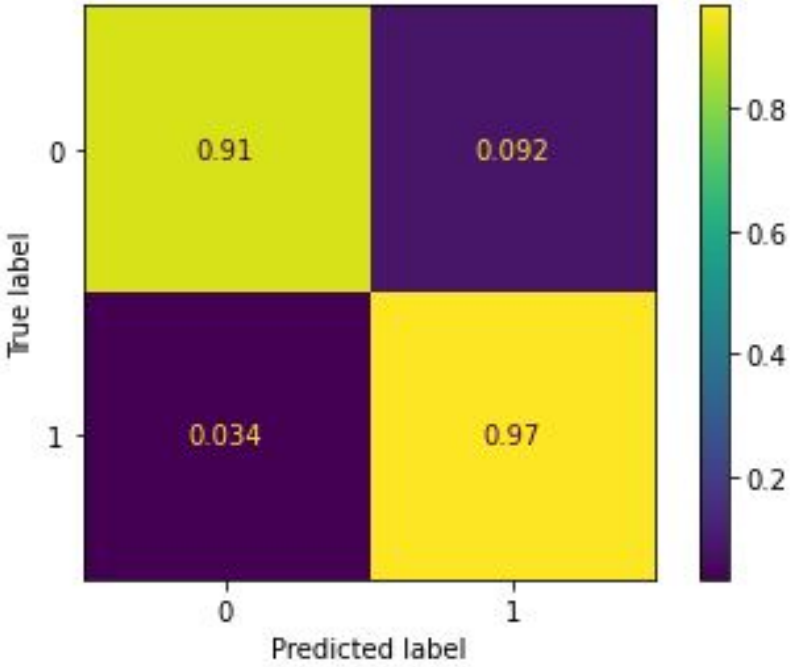
CLASSIFICATION AND EFFECTIVENESS

03

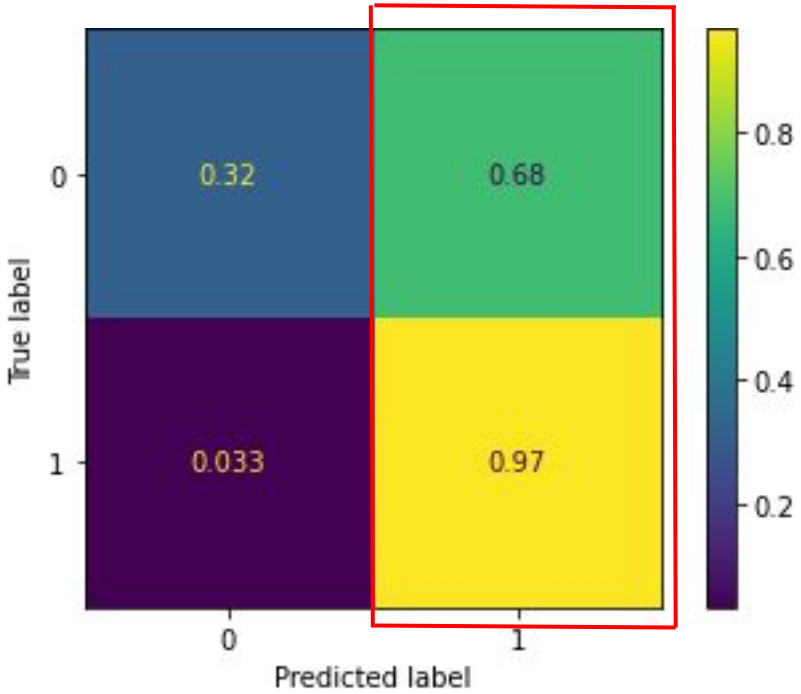
CLASS IMBALANCE

IDENTIFYING DRIVER ERROR

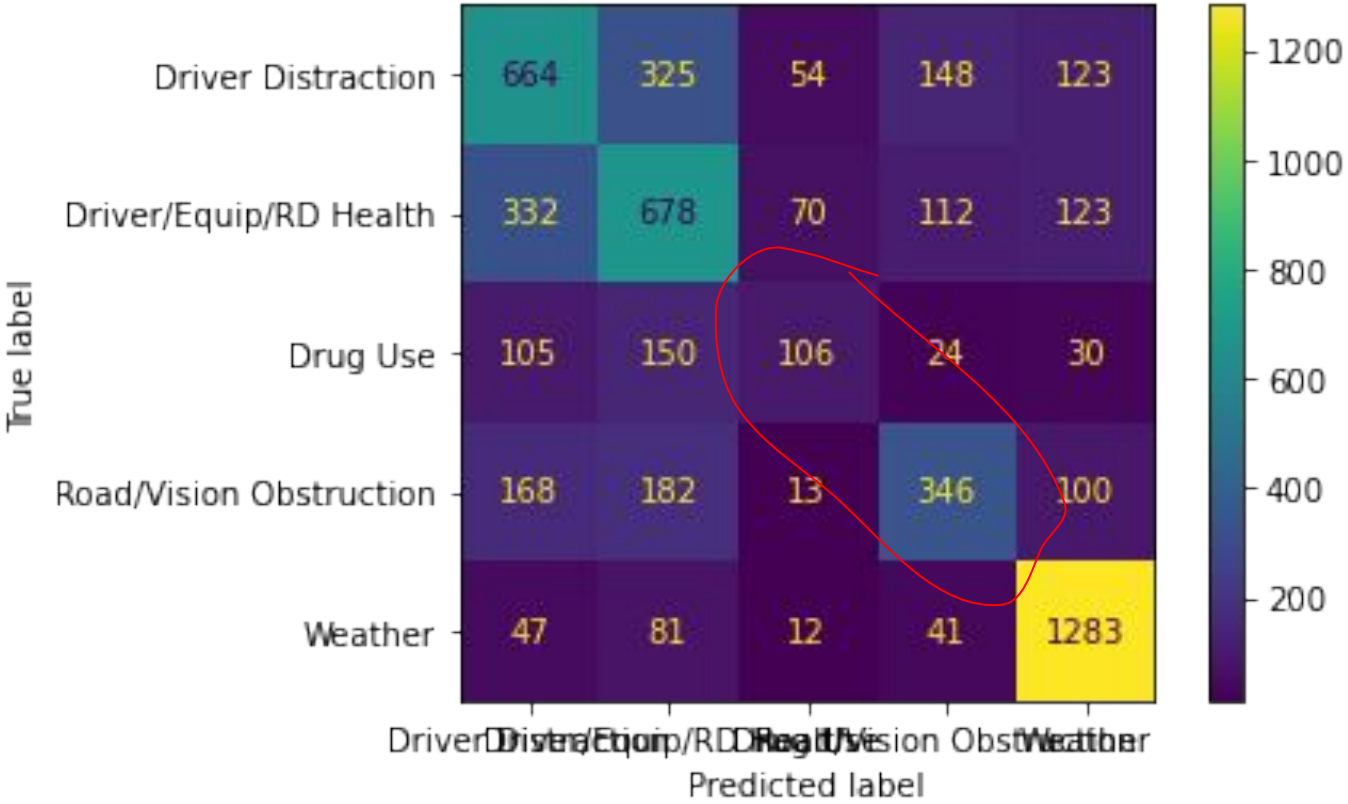
TRAIN RESULT

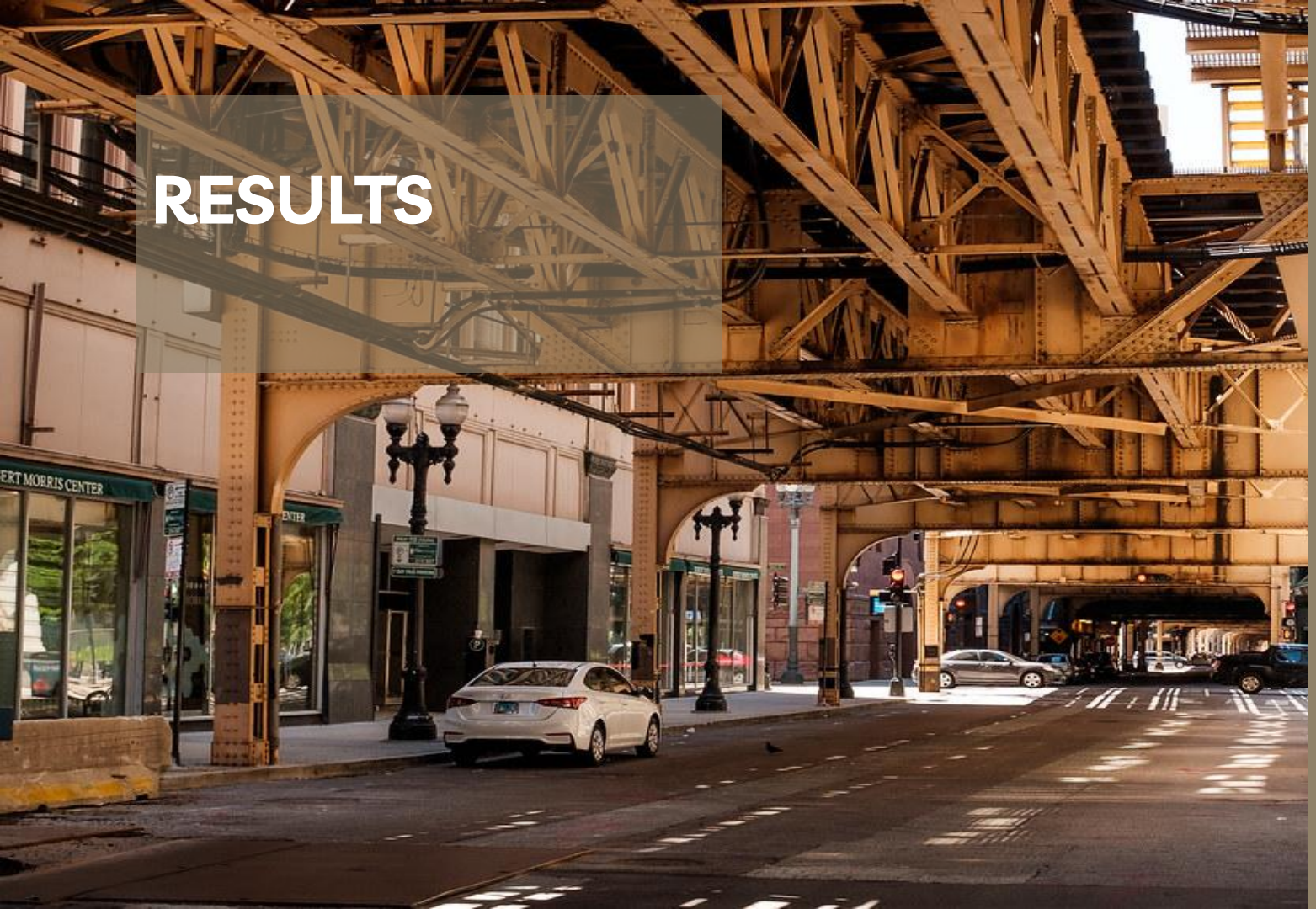


TEST RESULT



IDENTIFYING SUB-CLASSES





RESULTS

04

KEY FINDINGS

LOW INJURIES/FATALITIES

>99% were fatality free
>85% were injury free

TRAFFIC CONTROL

More injuries did occur WITH traffic control present
On average more accidents occurred WITHOUT traffic control

THE ROAD AHEAD

MARY'S MARK

05

NEXT STEPS

STEP 1

Improve data entry and collection

STEP 2

Clean an incorporate additional datasets for further insights
Identify problem traffic control areas

STEP 3

Model refinement and alternate classification groupings by association

THANK YOU
FOR YOUR
TIME