Collisions in Seattle:
Analysis & Prediction

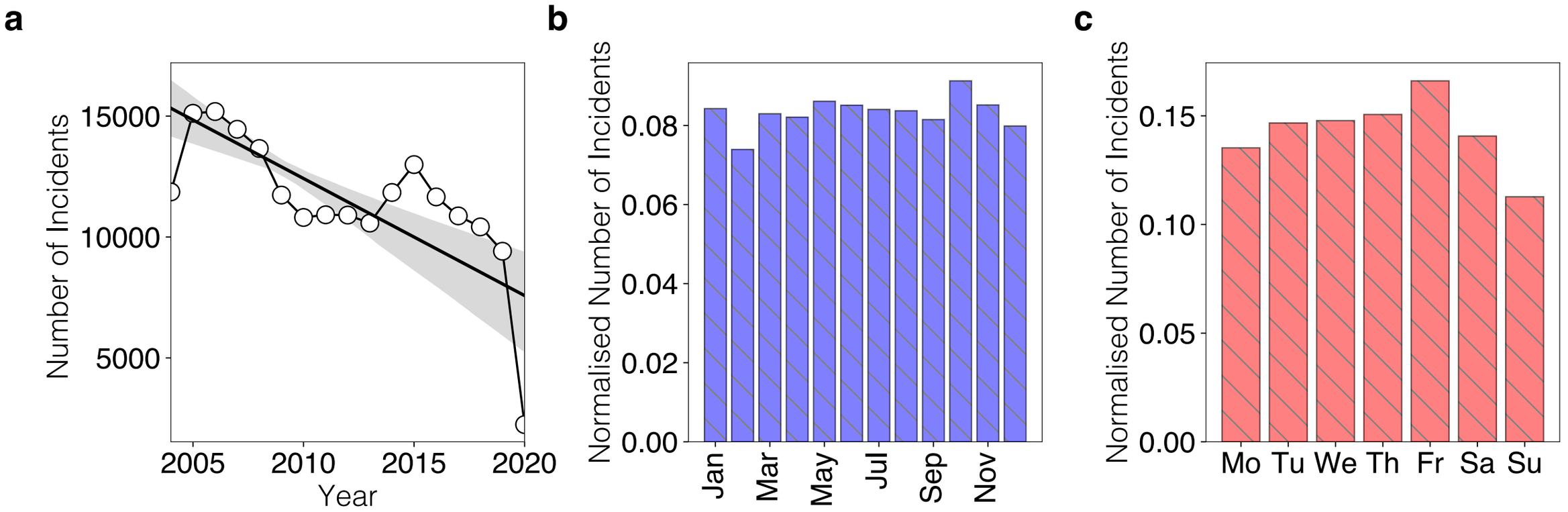
Mauleon-Amieva, A.

October, 2020

Collisions in Seattle

According to the Washington State Department of Transportation (WSDOT), car collisions occurs every 4 minutes in Seattle, MA.

Collisions in Seattle



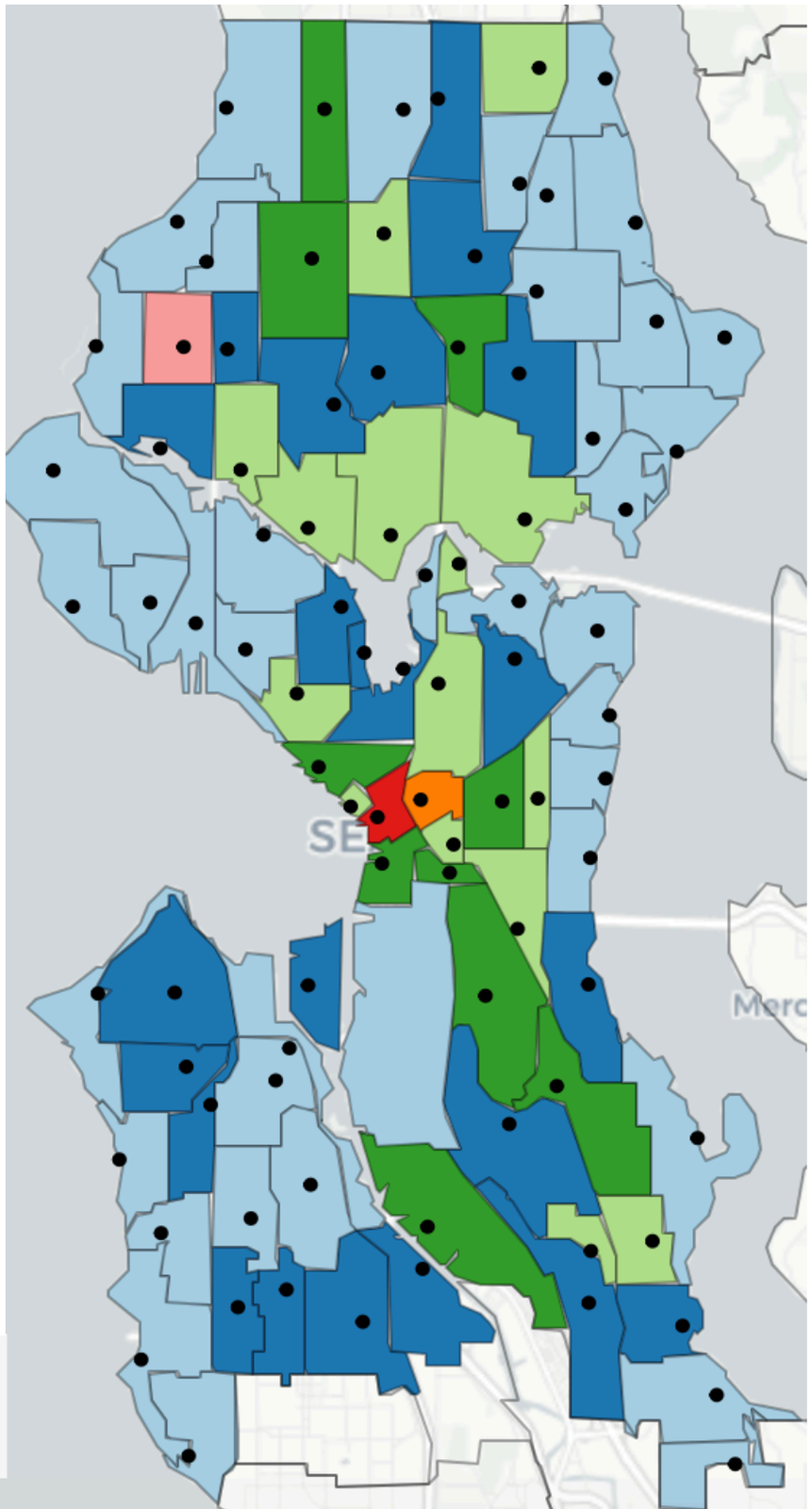
Historically, the number of collisions has decreased, as shown in Fig. **a**. Nonetheless, the number of incidents is evenly distributed across the year (Fig. **b**), and over weekdays (Fig. **c**)

Neighbourhoods

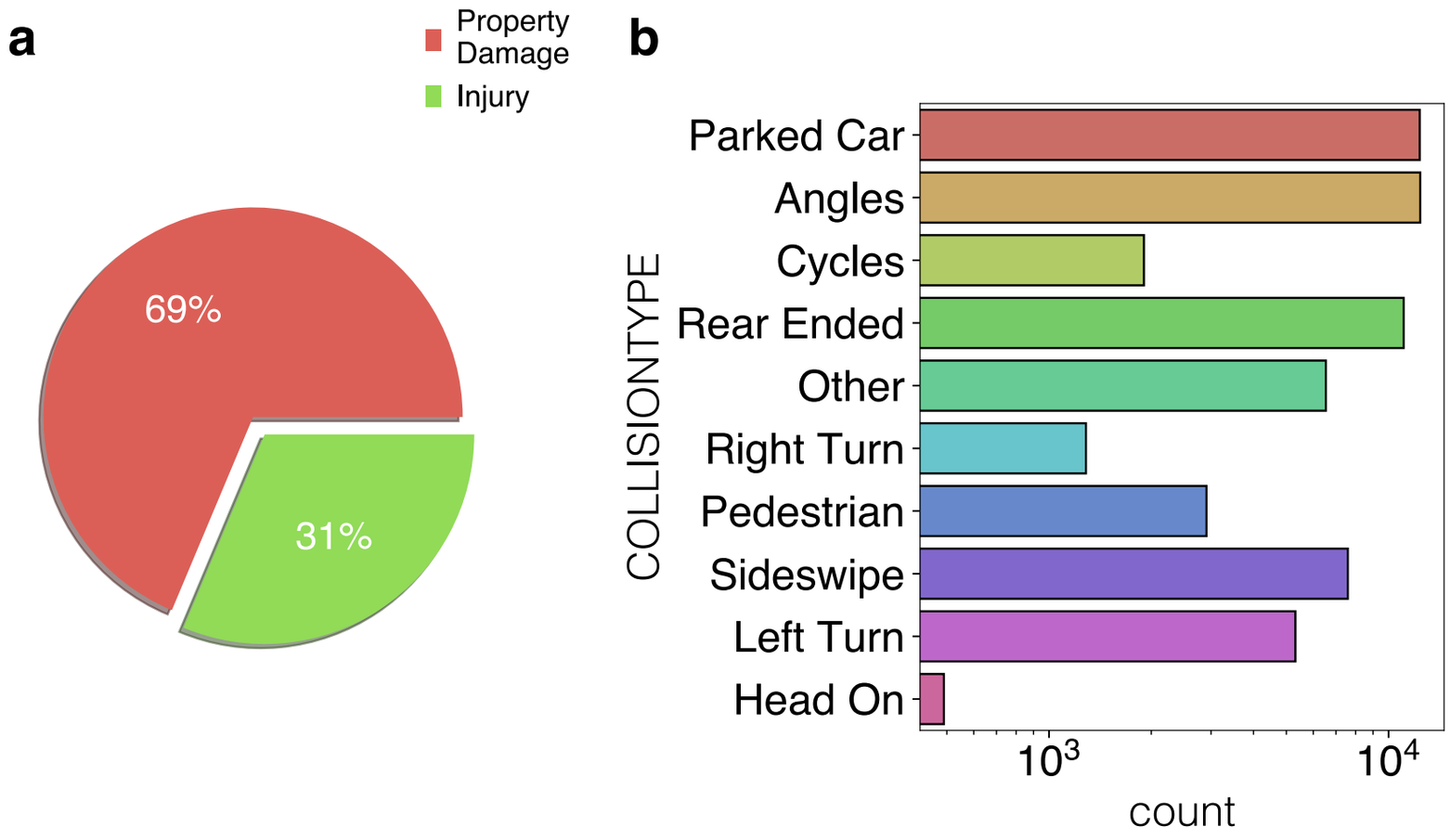
The number of collisions is not evenly distributed across the city.

Neighbourhoods with high rate (≥ 4,000 collisions) are identified.

First Hill and Central Business District are among the neighbourhoods with highest collision rate.



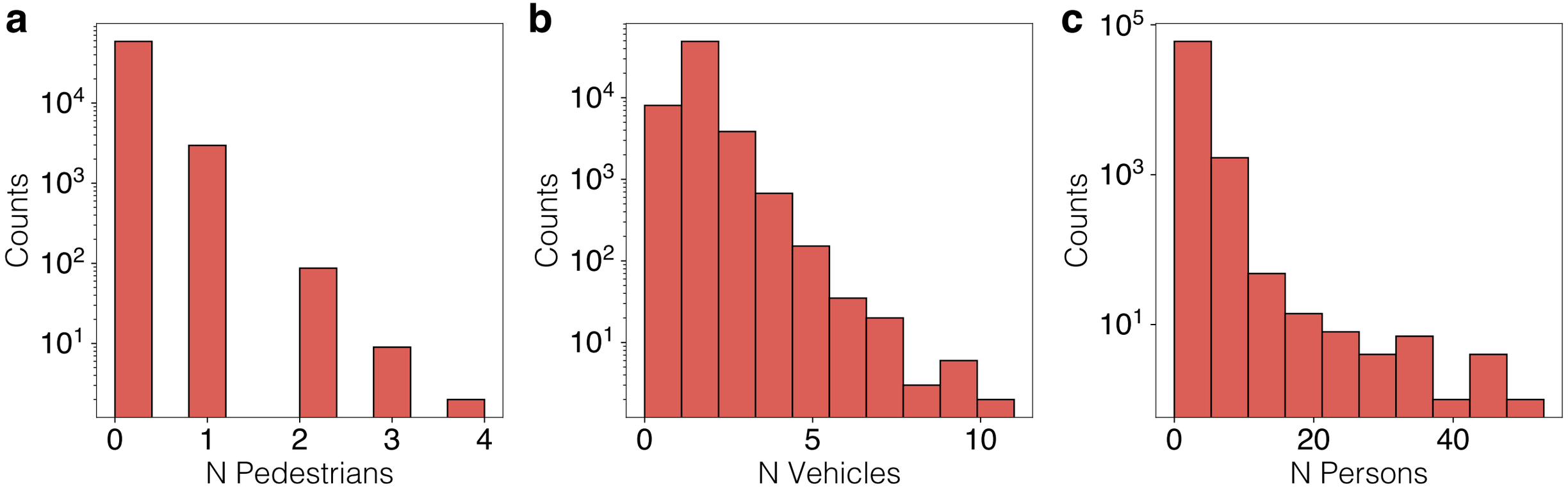
Type of Collisions



The severity and the type of collisions are investigated in the neighbourhoods of interest.

It is noted that most of the incidents result only in property damage (Fig. **a**). On the other hand, the type of accident is more varied. Most of the accidents involved parked cars and rear ended collisions (Fig. **b**)

Type of Collisions



Different types of collisions involved a different amount of persons and vehicles.

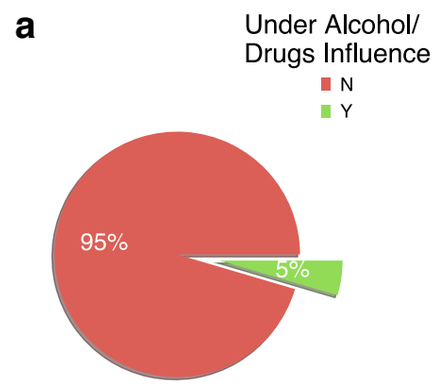
For the number of pedestrians, vehicles and persons, most of the collisions show a minimum number. However, few collisions involve more than 10 vehicles, and more than 40 persons.

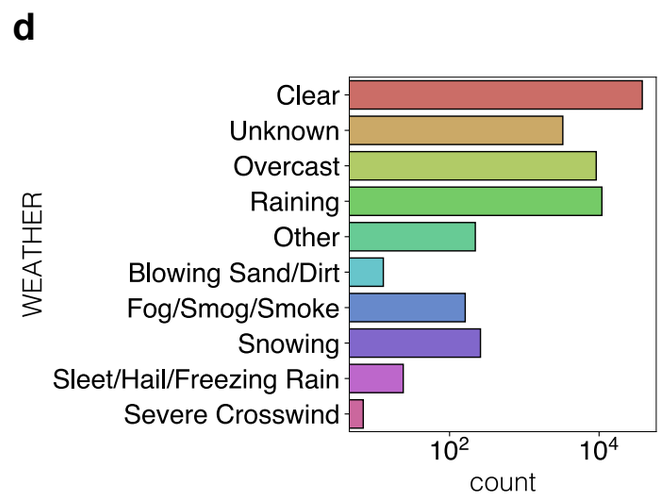
Conditions

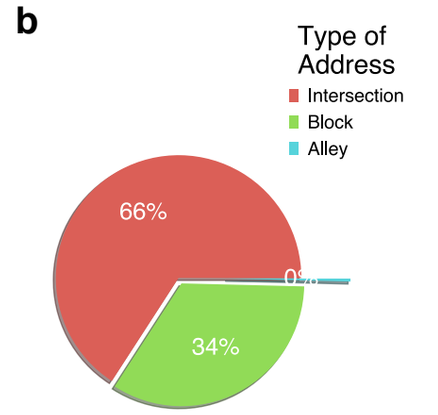
A series of different conditions are selected in order to extract information for the different types of collisions:

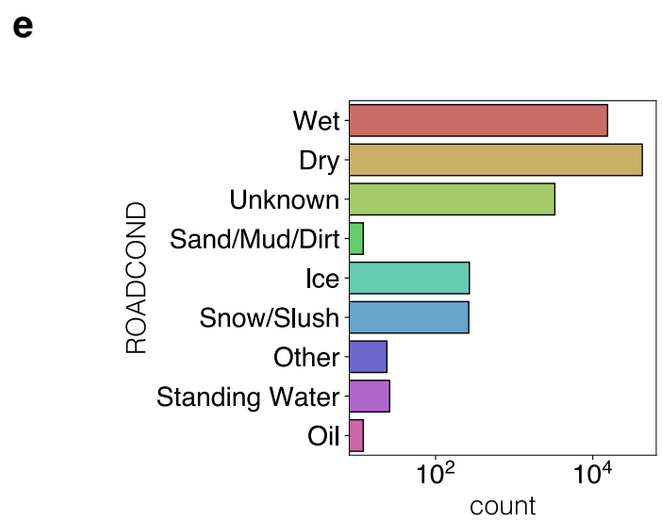
- Influence of alcohol/drugs
- Type of address
- Type of junction
- Weather
- Road conditions
- Light conditions

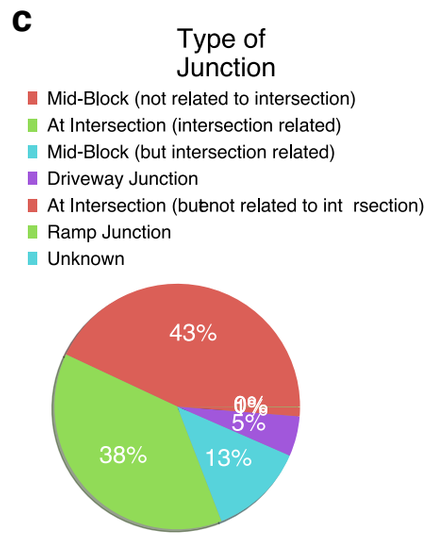
Interestingly, the majority of collisions occur within optimal conditions, e.g. clear weather, dry roads, and during daylight.

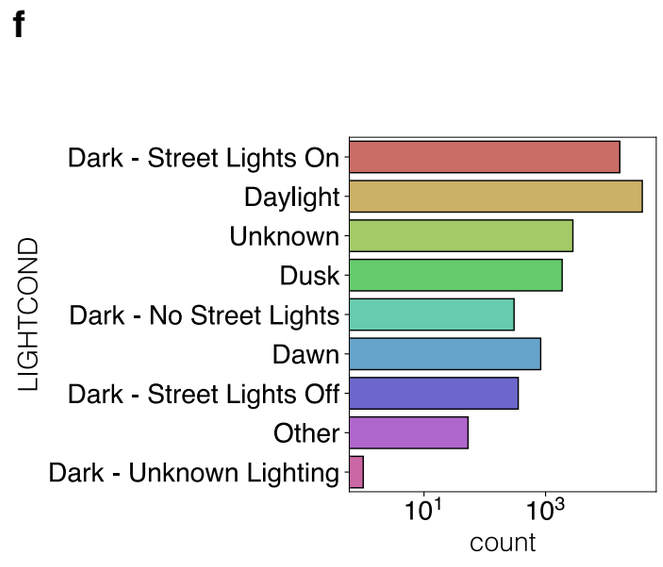




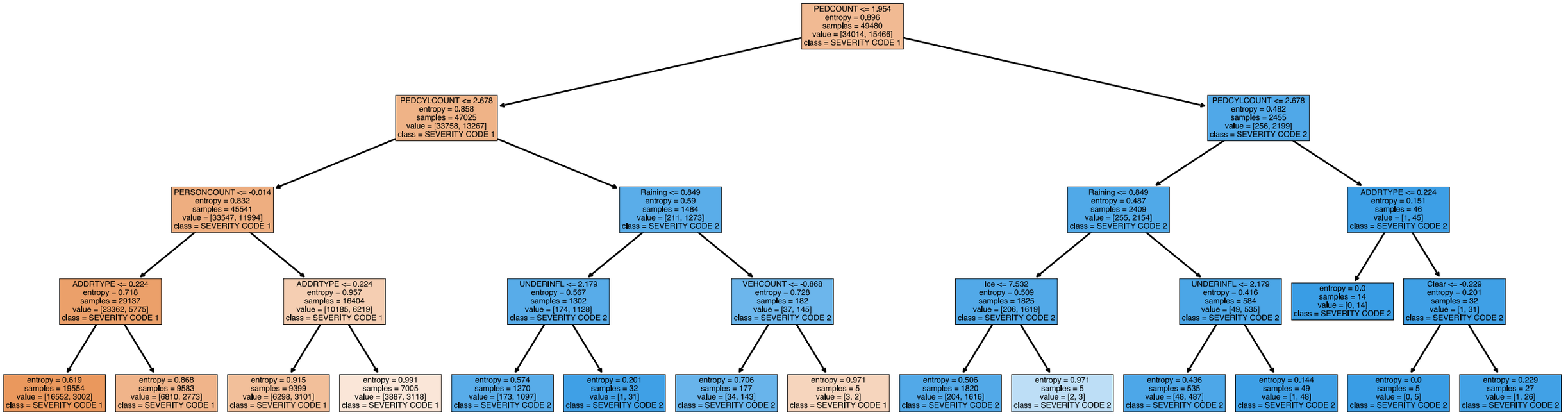








Modeling



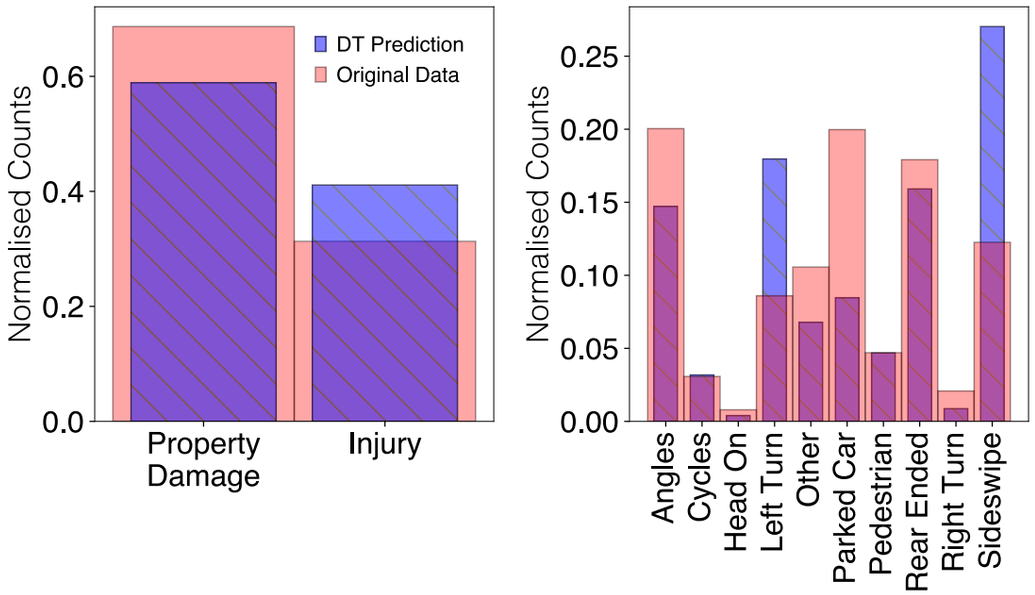
Different machine learning approaches are used to predict both the severity and the type of collision. The models used are:

- k Nearest Neighbours
- Decision Tree
- Support Vector Machine (SVM)
- Logistic Regression

Modeling

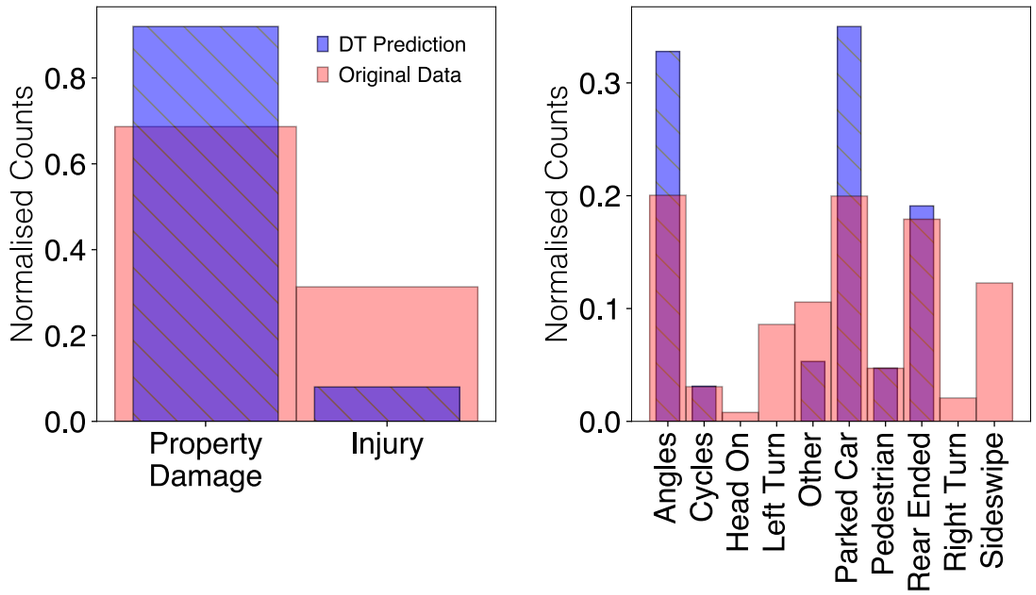
a

k Nearest Neighbours



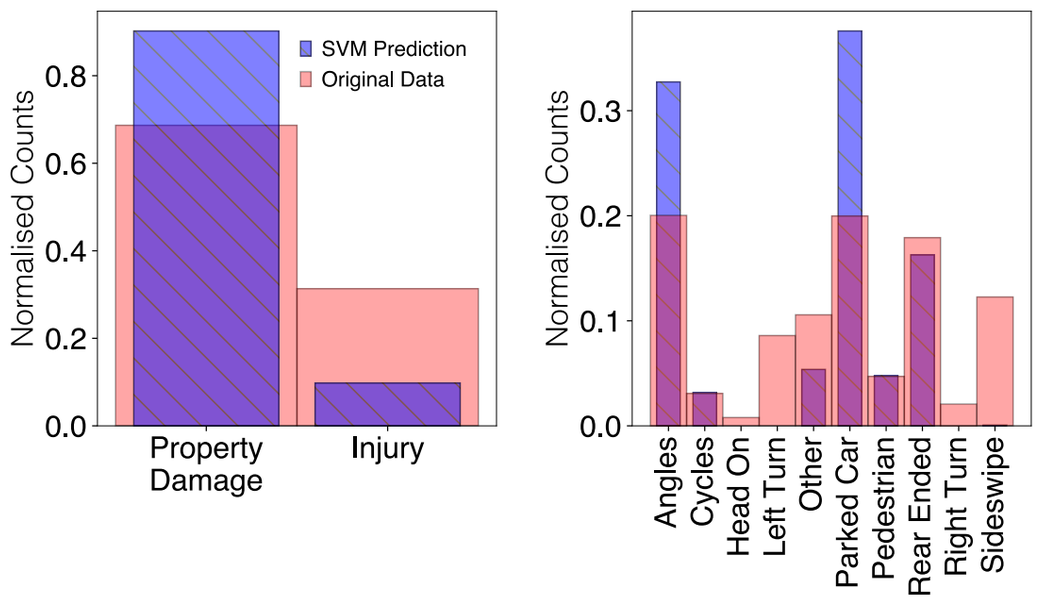
b

Decision Tree



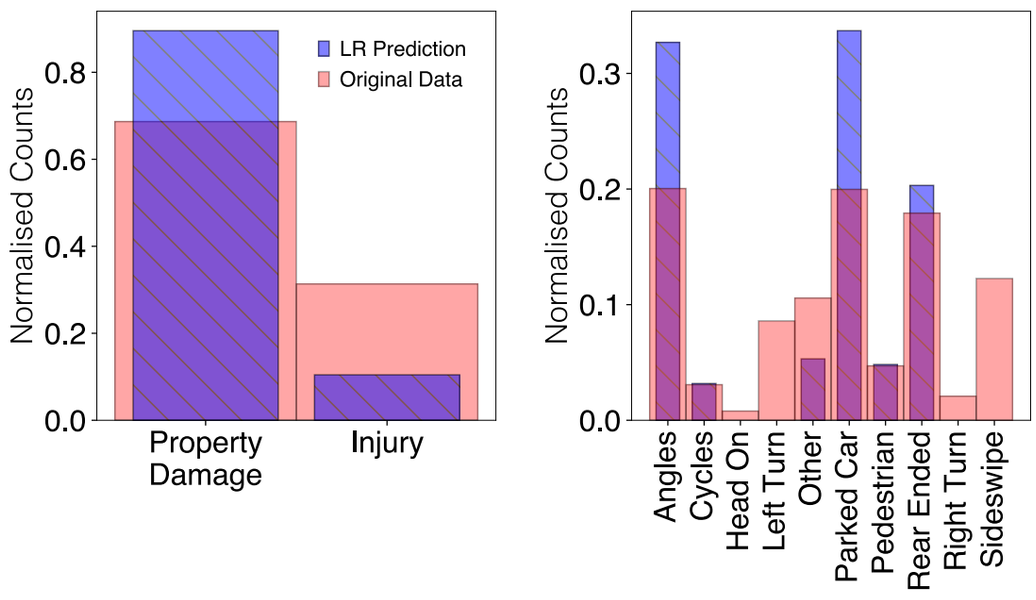
c

Support Vector Machine



d

Logistic Regression



Summary

The complete analysis for the modeling can be found here.

Overall, Logistic Regression resulted the best method to predict both the severity and the type on collision.