Exploration of Chicago Car Crash Data

Kevin Thompson, Dilruksha Jayaweera, Heber Nielsen, Sangrae Cho, Kumaraiah Pradeepkumar

Data Science Program, Southern Methodist University, USA

1   Introduction

Car accidents are no trivial matter in the United States. Roughly 800 car accidents occur every day in Illinois alone (Hilkevitch, 2013). Our project is to develop models to both predict and explain a subset of these entirely avoidable accidents in Chicago, Illinois. OpenWeather's historical weather database and the city government of Chicago's website to construct these models and insights. These datasets include factors that are widely believed to contribute to higher accident rates, such as weather, time, location, road conditions, and so forth.

The primary goal of our exploration is to explore useful features for predicting the number of accidents in a given police shift in location clusters called "police beats".

```python
import pandas as pd
import matplotlib.pyplot as plt

pd.options.display.width = 0
pd.set_option('precision', 2)

crashes = pd.read_csv("data/ChicagoCrashesPerHour2.csv")
print(crashes.describe())
```

##	SHIFT	BEAT_OF_OCCURRENCE	NUM_UNITS	INJURIES_TOTAL	INJURIES_FATAL	INJURIES_INCAPACITAT
## count	1.45e+07	1.45e+07	1.45e+07	1.45e+07	1.45e+07	1.45e
## mean	2.00e+00	1.21e+03	0.00e+00	0.00e+00	0.00e+00	0.00e
## std	8.16e-01	6.99e+02	0.00e+00	0.00e+00	0.00e+00	0.00e
## min	1.00e+00	1.11e+02	0.00e+00	0.00e+00	0.00e+00	0.00e
## 25%	1.00e+00	6.32e+02	0.00e+00	0.00e+00	0.00e+00	0.00e
## 50%	2.00e+00	1.12e+03	0.00e+00	0.00e+00	0.00e+00	0.00e
## 75%	3.00e+00	1.81e+03	0.00e+00	0.00e+00	0.00e+00	0.00e
## max	3.00e+00	2.54e+03	0.00e+00	0.00e+00	0.00e+00	0.00e

References

Hilkevitch, J. (2013). Traffic accidents on the rise in state, to almost 800 a day. *Chicago Tribune*.