CAR ACCIDENTS PREDICTION

IN CHICAGO

FEATURE ENGINEERING

Objective

Car accidents number prediction based on weather condition data

Why Chicago?

- **Big traffic**
- **Changeable weather**

Business Value

- **Proactive Police Arrangement**
- **Car Services**
- **Safe Routing**

DATA

Sources: Weather underground
Chicago Data Portal

Potential significant features:

NUMERICAL

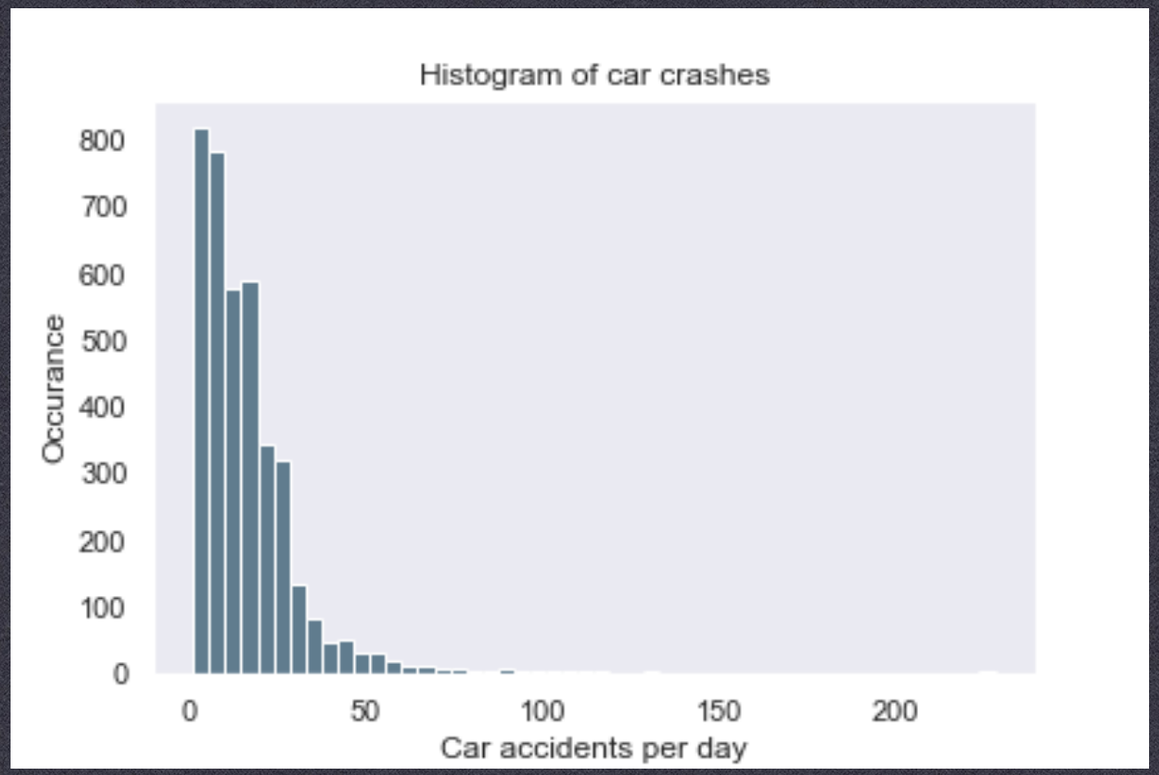
- Humidity
- Wind Speed
- Pressure
- Temperature
- Dew point

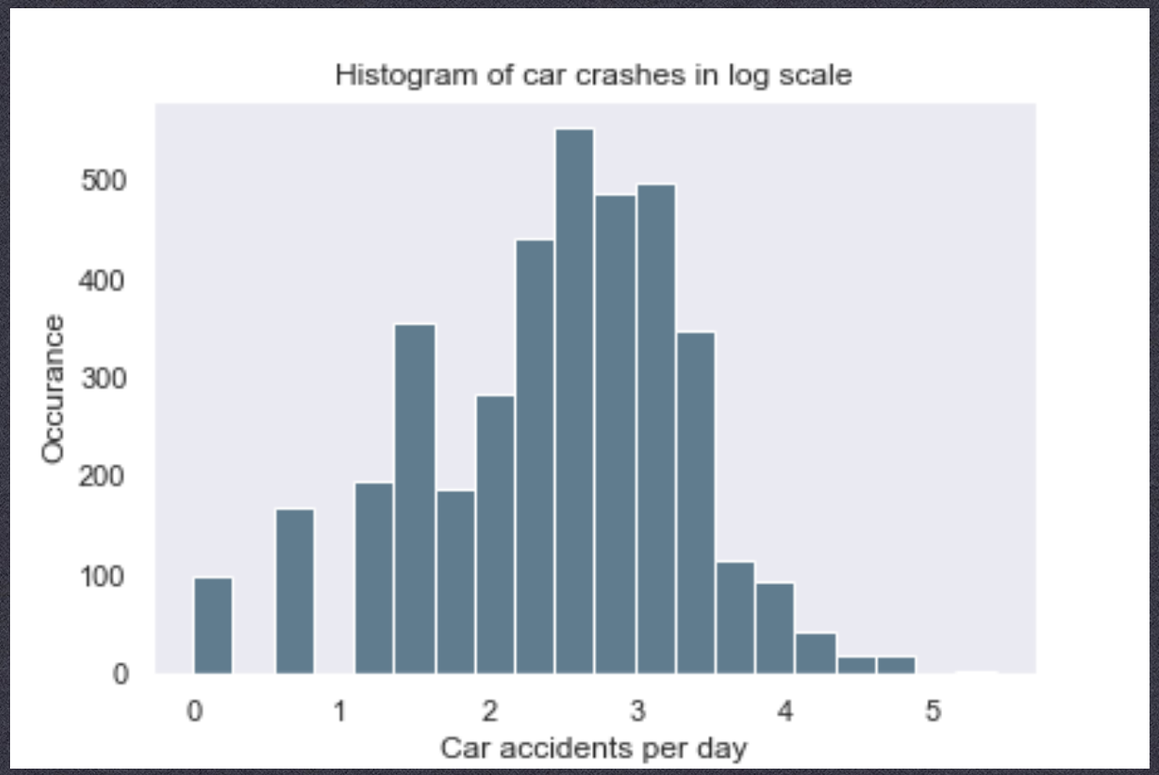
CATECORICAL

- Weather condition:
  Clear Rain Snow
- Lighting condition:
  Dark Light

DATA MANIPULATION

DROPPING NANS
REMOVING OUTLIERS
SCALING
TRANSFORMATION IN LOG SCALE
OVERCOME MULTICOLLINEARITY





FEATURE ENGINEERING

? Dark time + Rain

? Dark time + Snow

? Month, Day of Week, Hour

? Polynomial features

? Intersections

DOES THE REGULARIZATION MATTER? A LOT!

	Linear Regression without Regularisation	Linear Regression with Lasso Regularisation
CROSS-VALIDATION R^2 adj		
Mean	-1.34e+20	0.5714
Variance	4.02e+20	0.0556
Hold-Out Errors		
MAE	56e+9	0.9110

FURTHER WORK

(TO INCLUDE OTHERS FEATURES NOT CONNECTED WITH WEATHER CONDITION FEATURES)

SPEED LIMIT

40, 50, 60

TRAFFIC TYPE

Divided, One-way

TRAFFIC CONTROL DEVICE

Stop signs, Stop signals