Chicago Car Crash Prediction
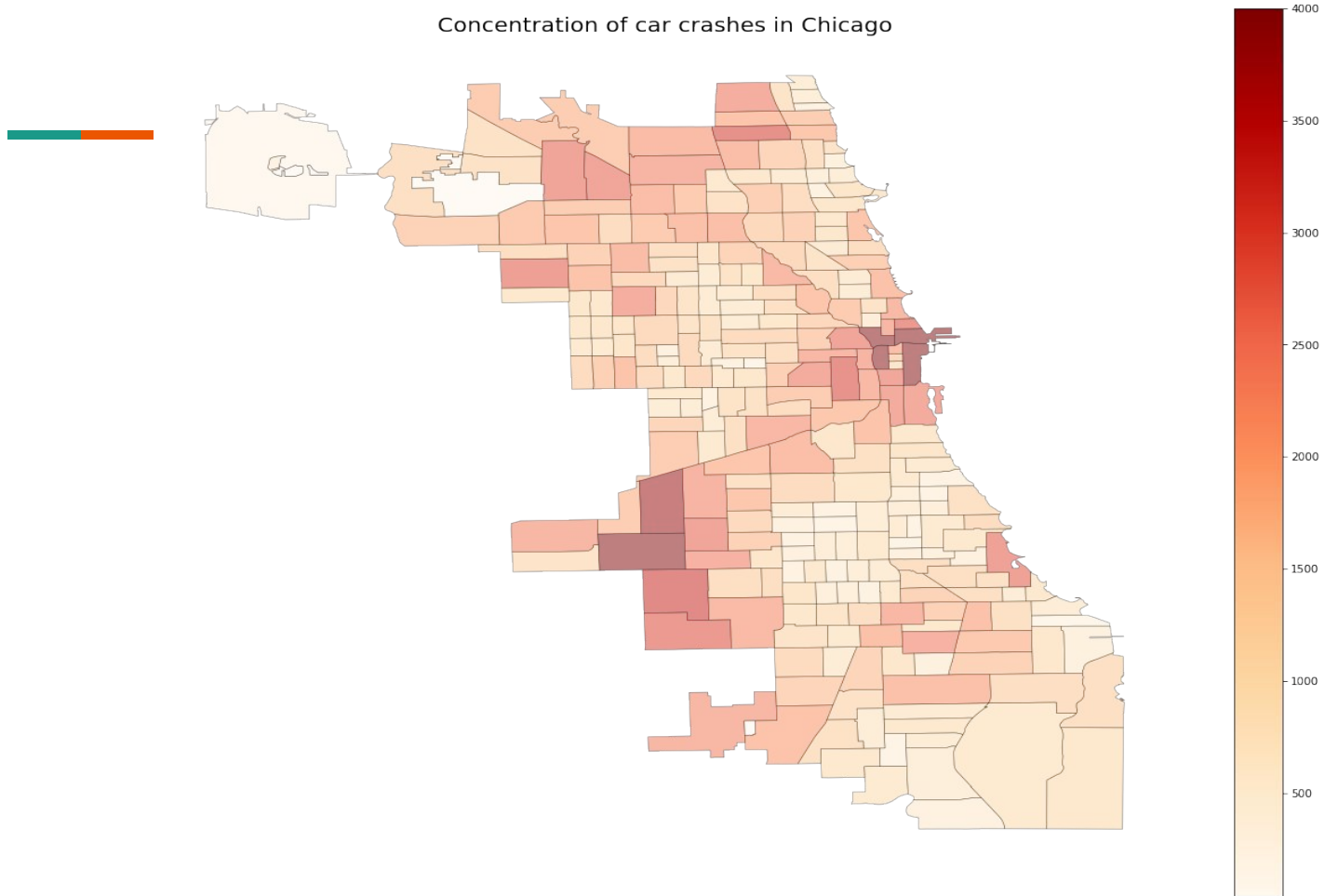
by Nino Delgado and Vittorio Scacchetti
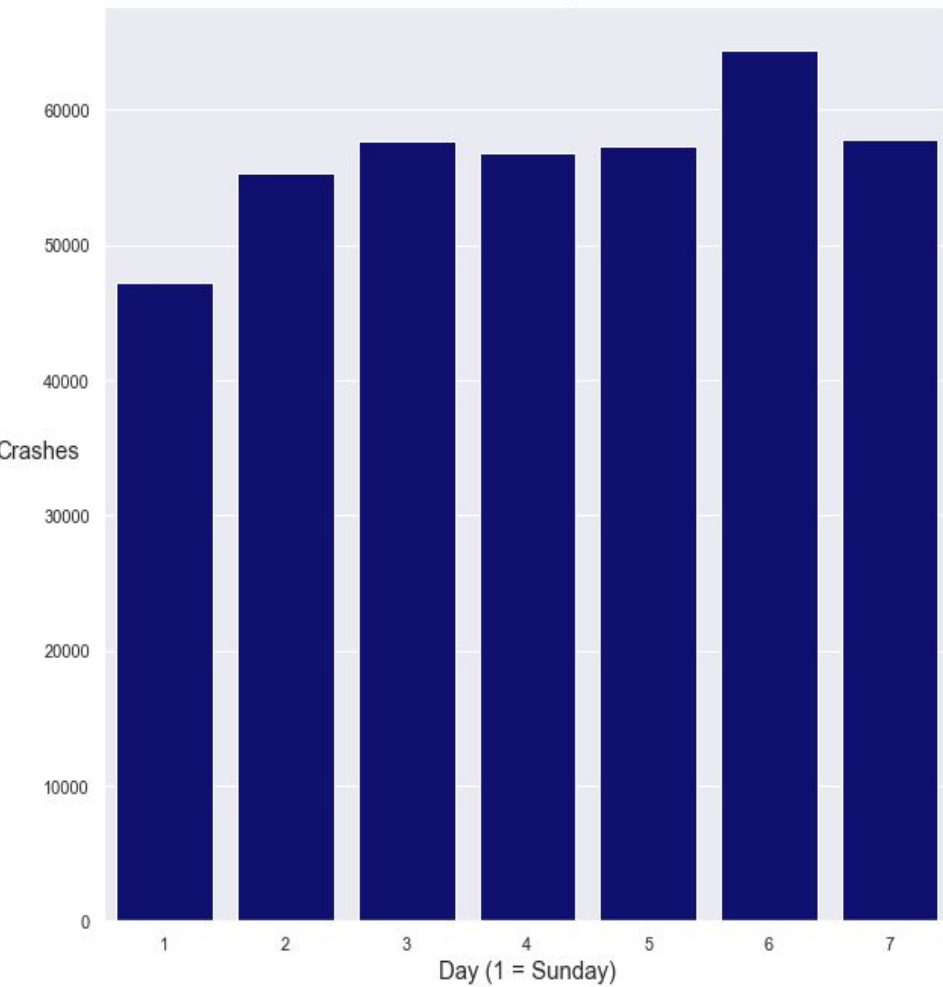
Objective and data

The aim of the project is to determine the major cause and costs of traffic crashes in Chicago, IL.

The data was provided by the City Of Chicago Data Portal, which contained information regarding the crashes and the people and vehicles involved in the crashes.
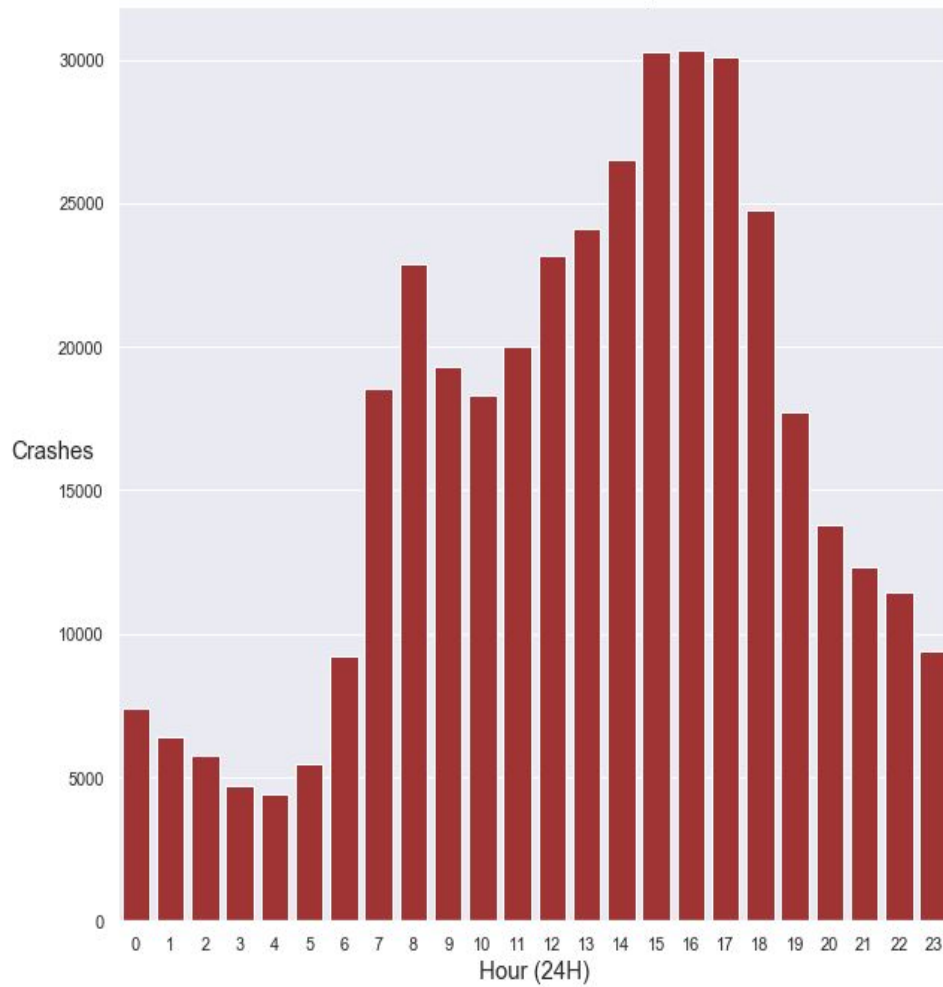
Concentration of car crashes in Chicago

Crash based on the day of the week

Crash based on the hour of day

Classifier: Decision Tree

train dataset: 0.8961

validation dataset: 0.8974

Classifier: Bagging

train dataset: 0.8986

validation dataset is: 0.8999

Classifier: Random Forest

train dataset: 0.95

validation dataset: 0.892

Preliminary advice:

- improve viability during working hours
- increase safety and control measures in highly trafficked areas

Future improvements:

- Selecting features
- Calculating confusion matrices

Thanks for listening!