Chicago Traffic Accidents

Determining the Primary Causes Of Crashes from 2015 to 2024

Joel Mott

Flatiron School

May 2024

business context

stakeholder:

Chicago Department of Transportation

goal:

reduce traffic crashes by determining their main causes

project requirements:

keep project methods *interpretable*

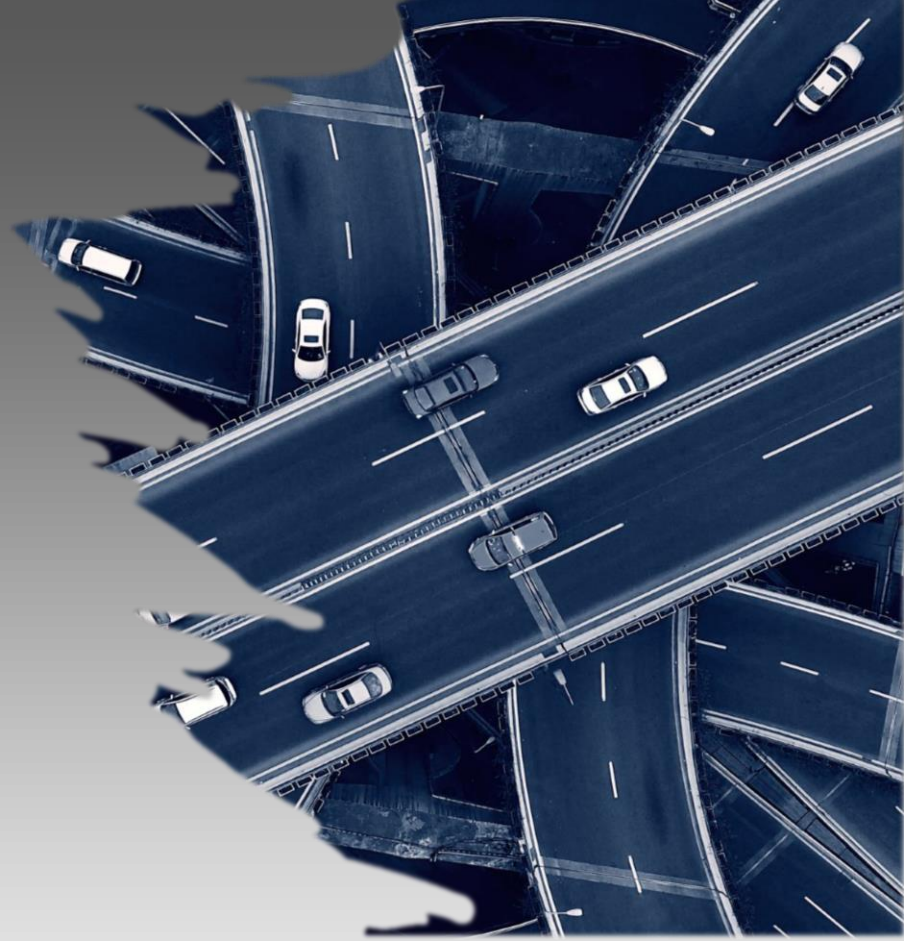
project overview

perspective: a broader look at safety

target: whether an accident was avoidable

business problem: reducing complex data to an interpretable prediction process

original dataset



- from the City of Chicago website
- detailed records dating back to 2015
- three separate files: crashes, people, and vehicles
- when combined: 3.8 million rows with 146 columns

Methodology

narrow columns down

try different modeling techniques

optimize the best model

reducing columns

146 columns = too many for an **interpretable** model

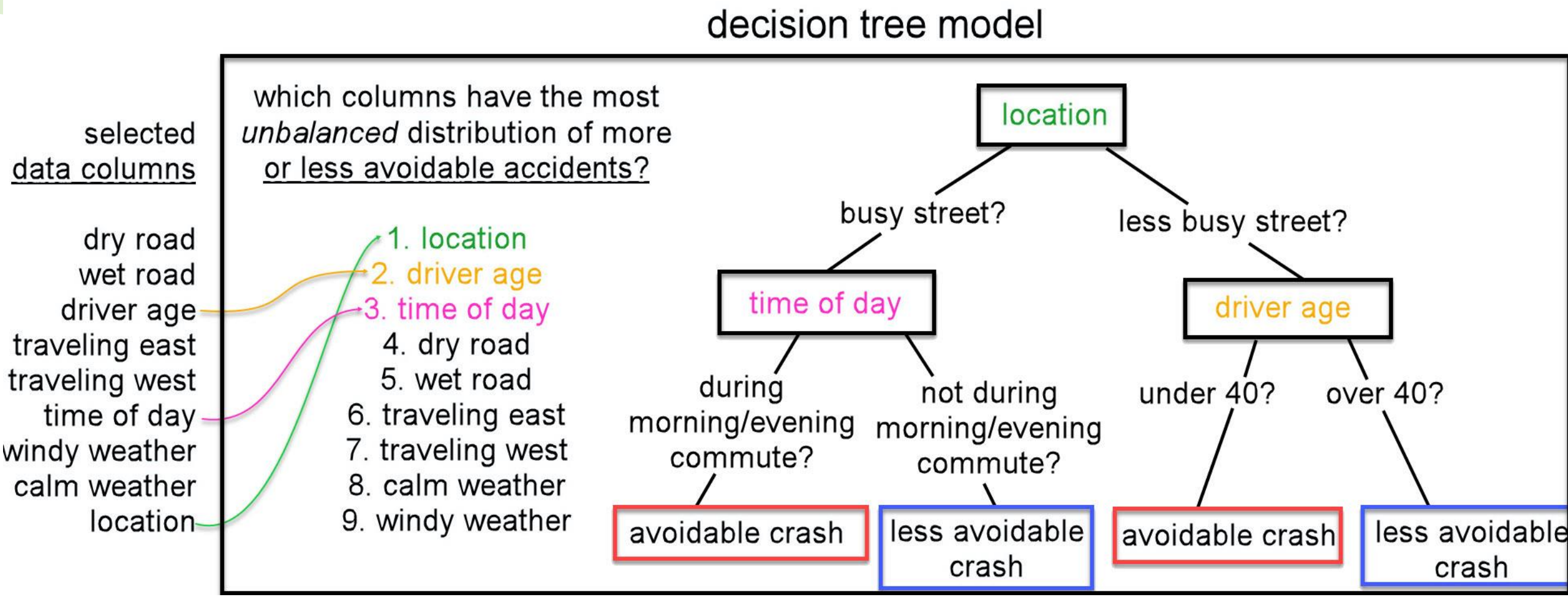
unnecessary columns reduced via:

- domain knowledge
- inherent redundancy

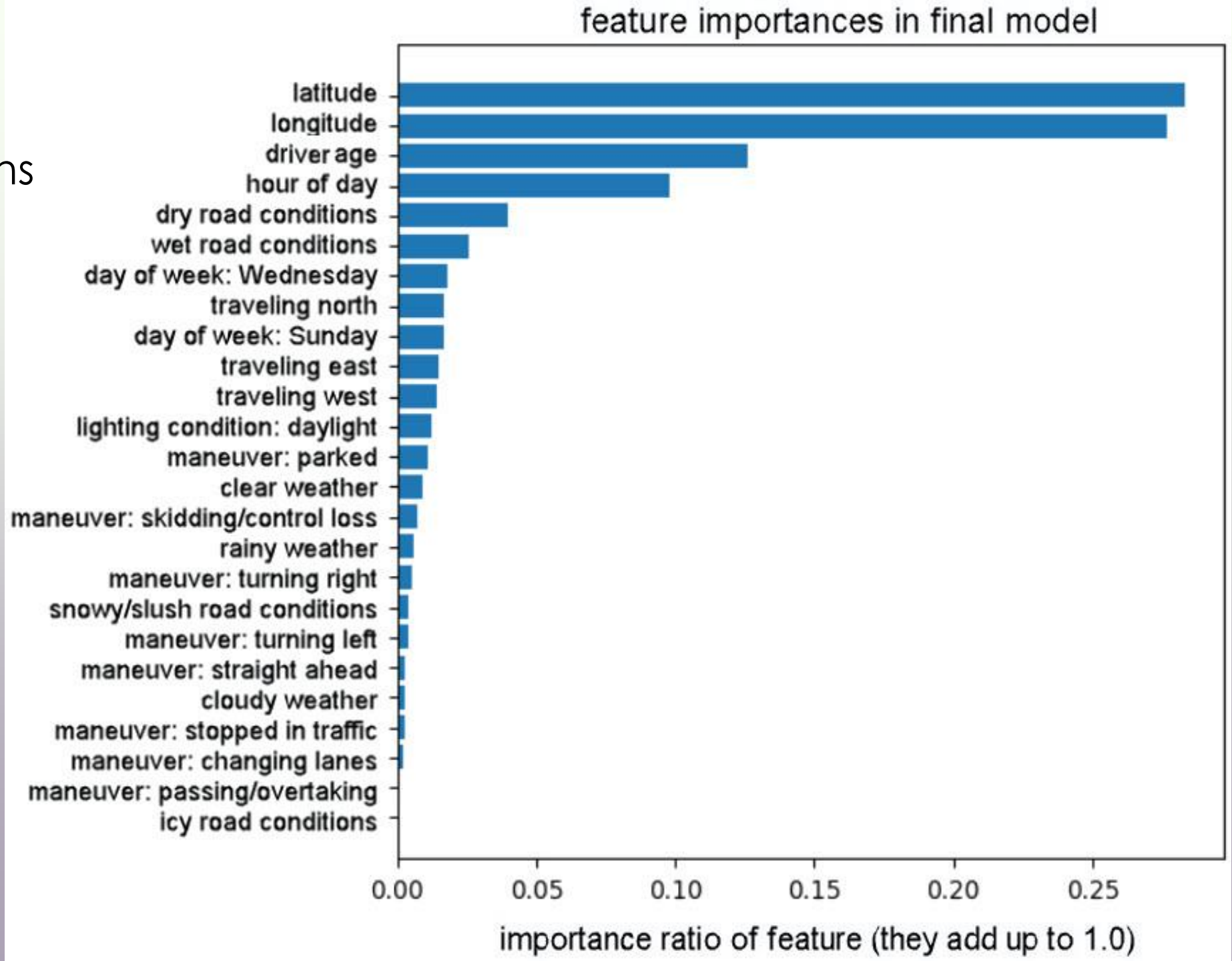
further columns reduced by:

- models that tell **how much data is explained** by each column
- **statistical** models

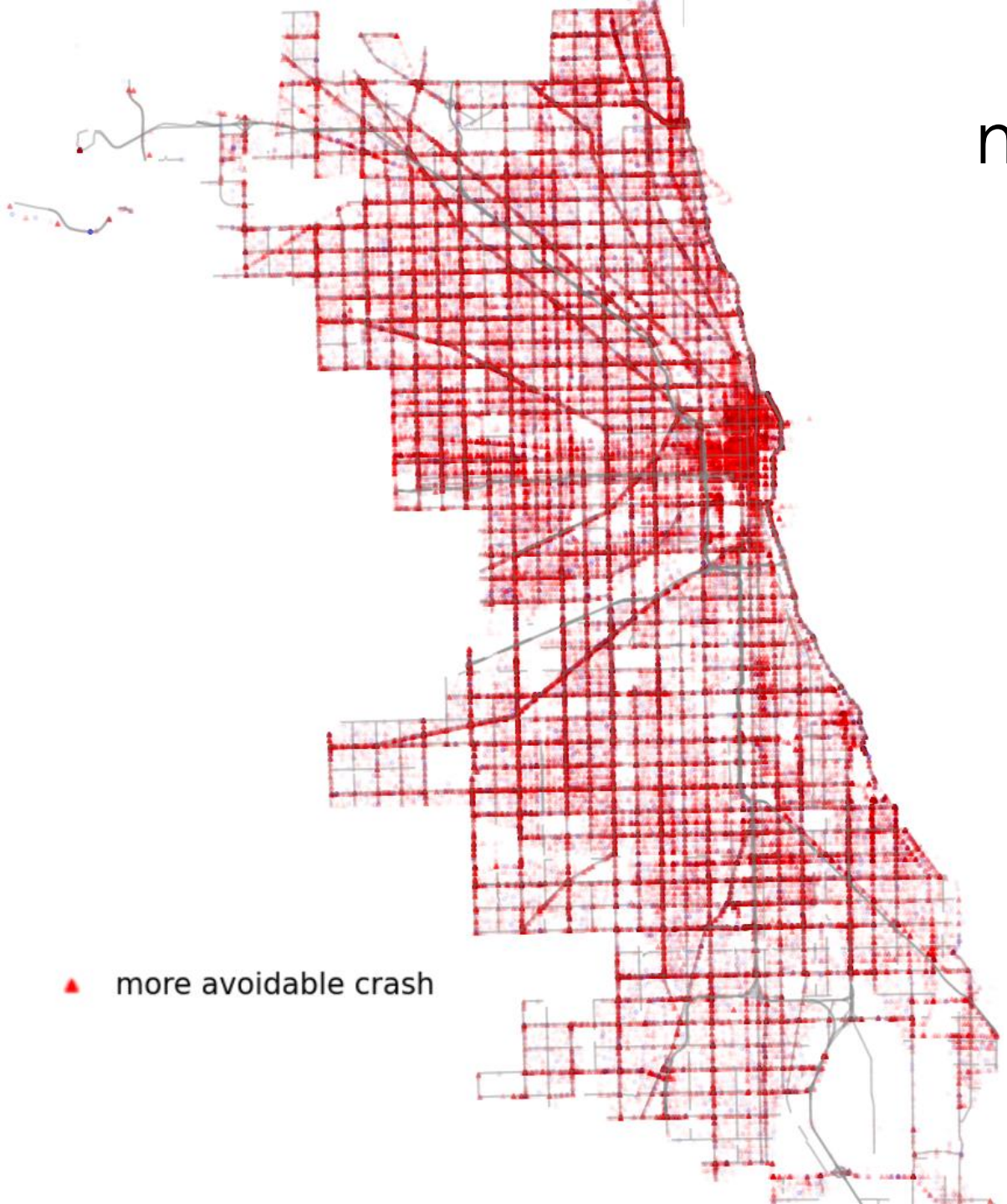
simplified decision tree example



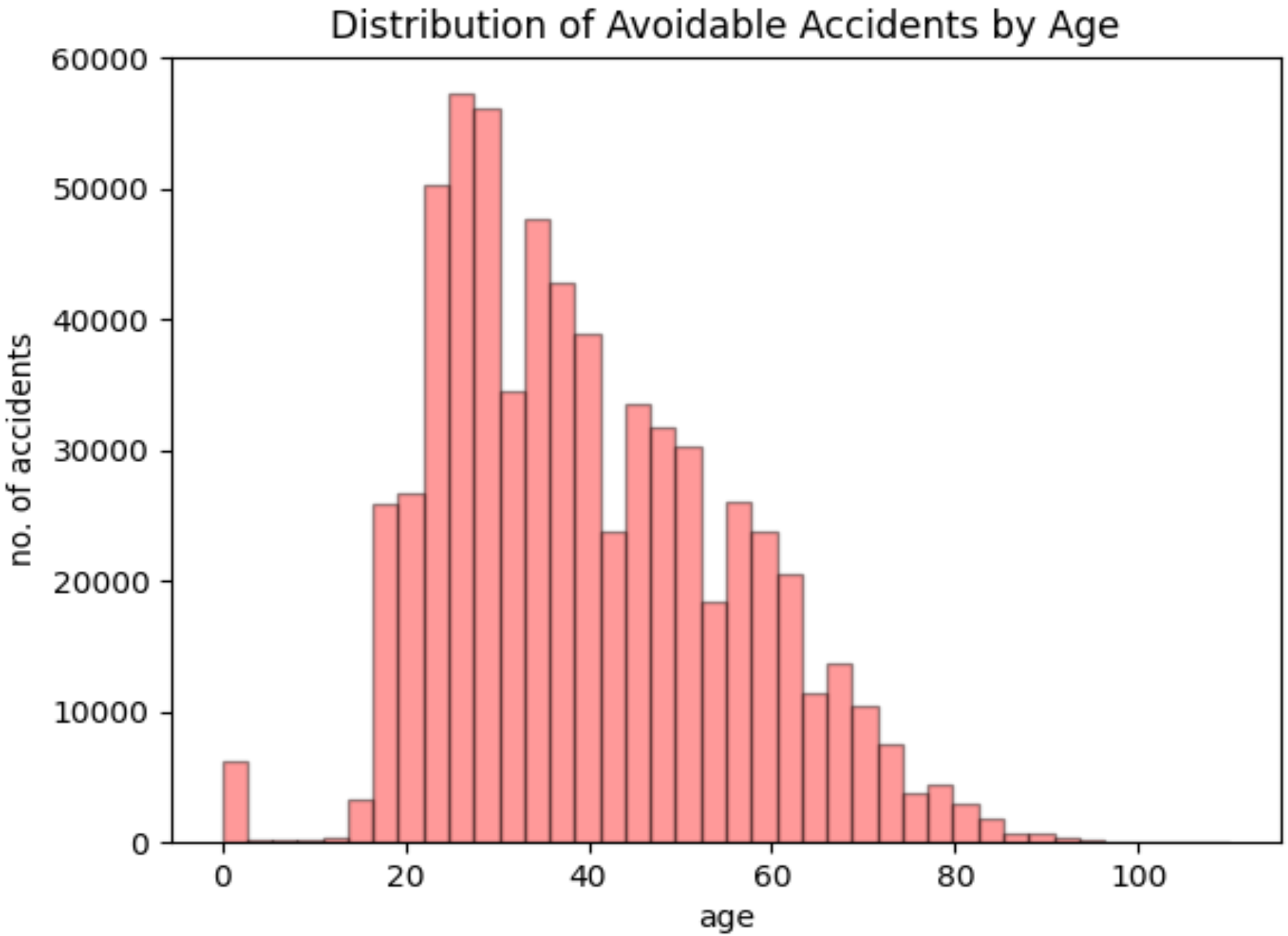
Model results:
most important columns



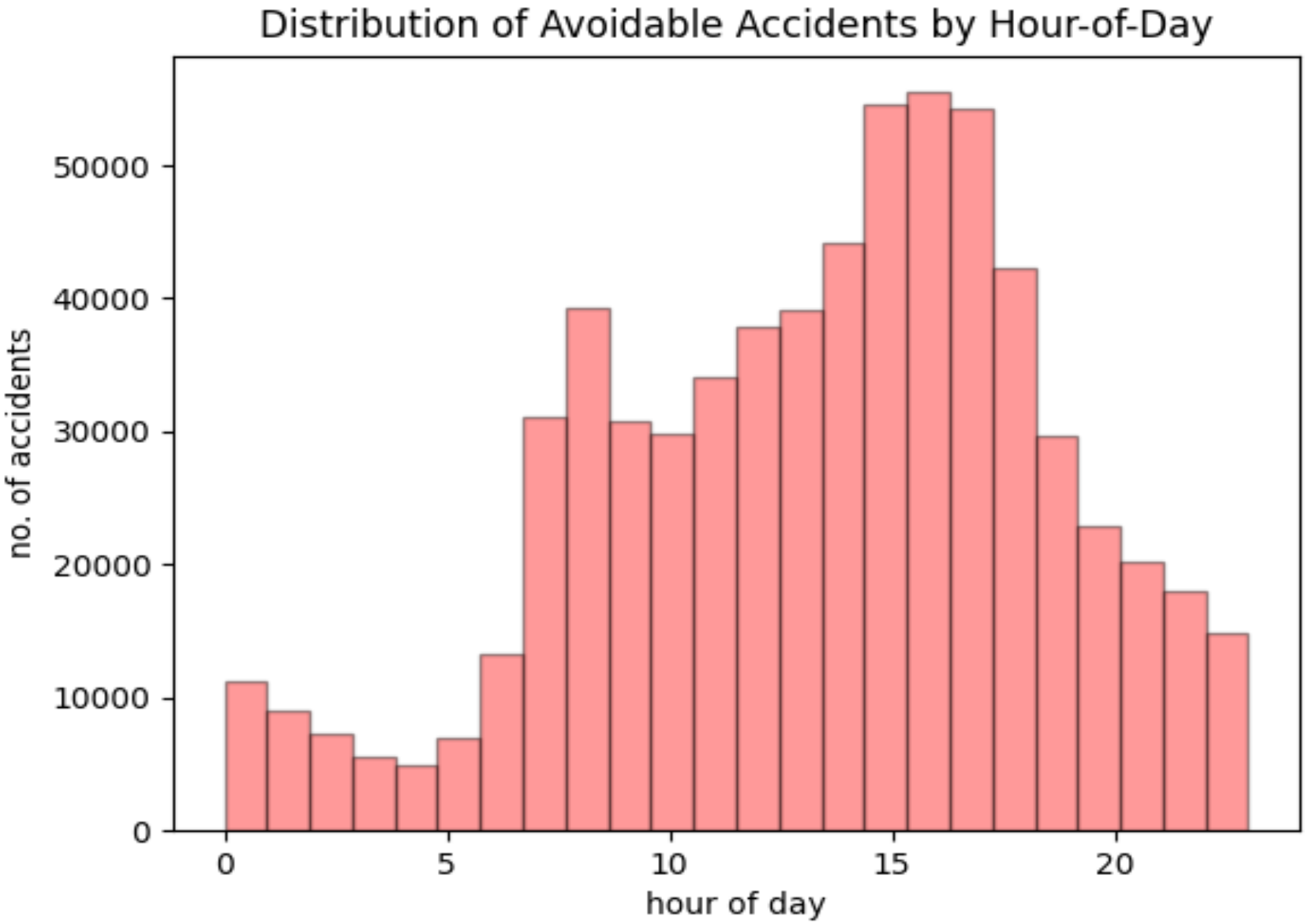
mapping more avoidable accidents



targeting drivers aged 23-40



commutes see the most crashes



recommendations

Ads focused on safety in **heavy traffic** for **drivers between 23-40**

Road sign/traffic signal studies in Chicago's **middle/downtown**

Safety PSAs over the radio **during commute times**


Thank You
joel.mott8@gmail.com
Flatiron School