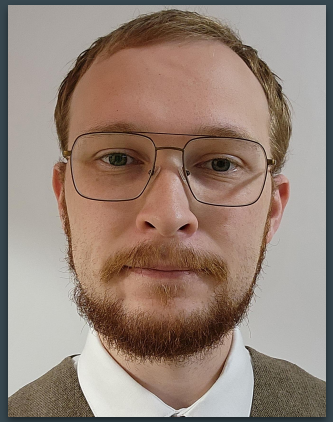
Analysis of Traffic Accidents In Chicago

November 17, 2023



John Baumgartner

Technical Lead

Github: jbaumgit

Email: jtbpilgrim@gmail.com

Dan Rosen

Github/Presentation Lead

Github: dangrosen

Email: dan_rosen@outlook.com

Business Problem

Data Overview

Analysis

Recommendations

Future Insights

Disclaimer: there are images of car crashes throughout the presentation

Findings

Accidents in the evening and overnight are about 8% more likely to cause damage over $1500, compared to other times of day.

Business Problem

Business Problem

- Private passenger vehicle crash in Chicago
- Property damage over $1,500
- Meaningful predictions?

Data Overview

Data Overview

Data

- 766,595 records on car accidents
- Data from the Chicago Data Portal

Conditions

- Crashes from 2021-2023
- Drivers of Private Passenger Vehicles
- 326,549 records remaining

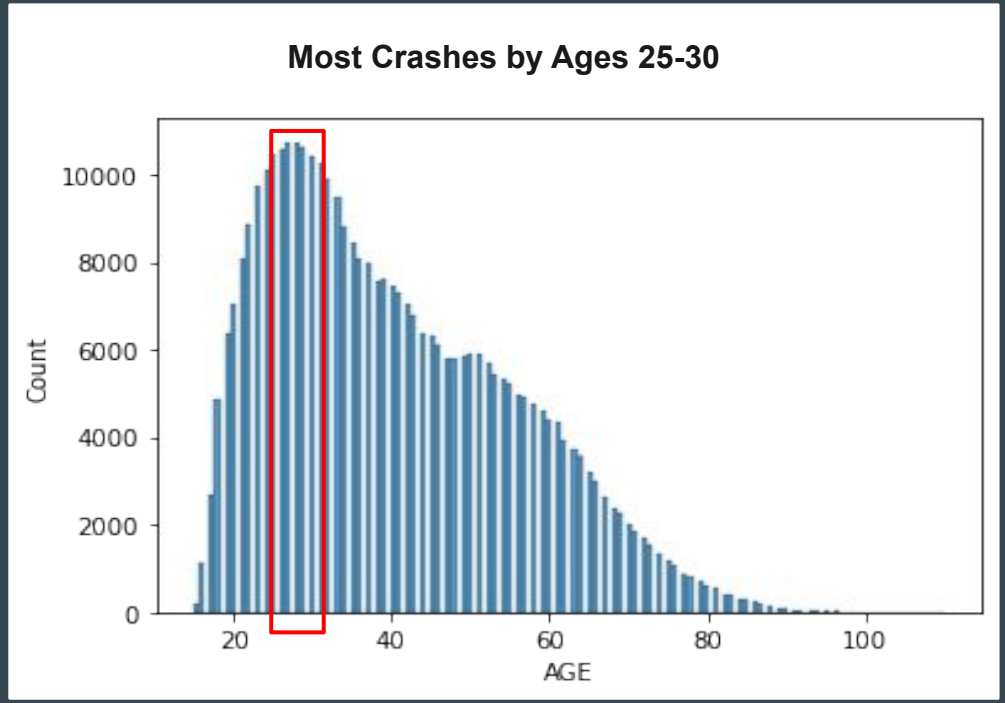
Limitations

- ~50% self reports
- Damage estimated
- Culpability not included

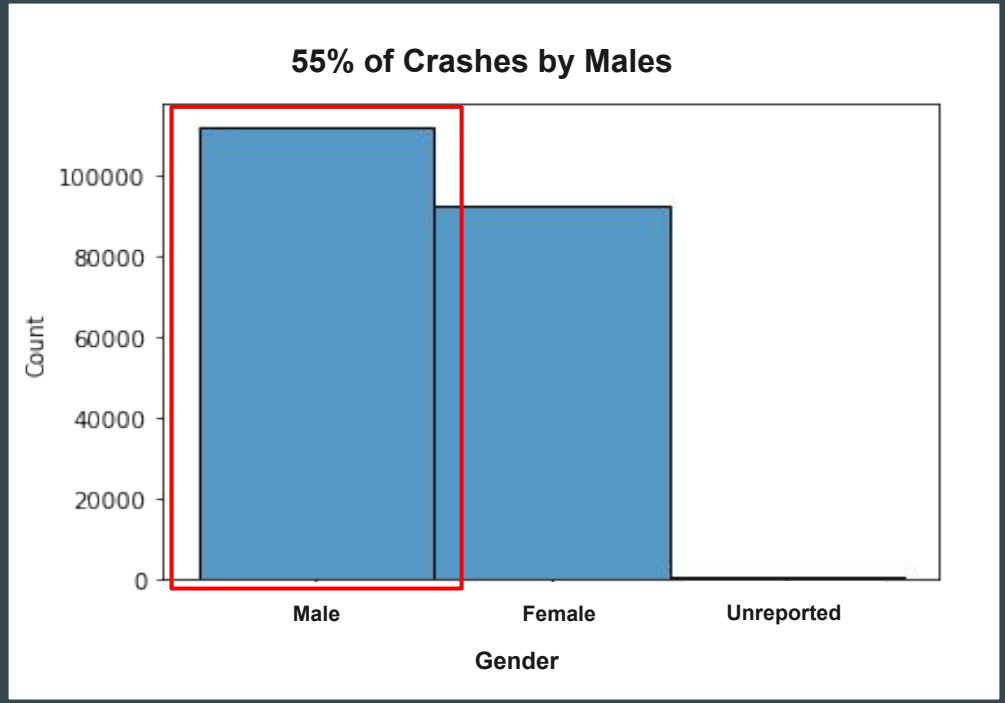
Analysis


Business Problem
Data Overview
Analysis
Recommendations
Future Insights

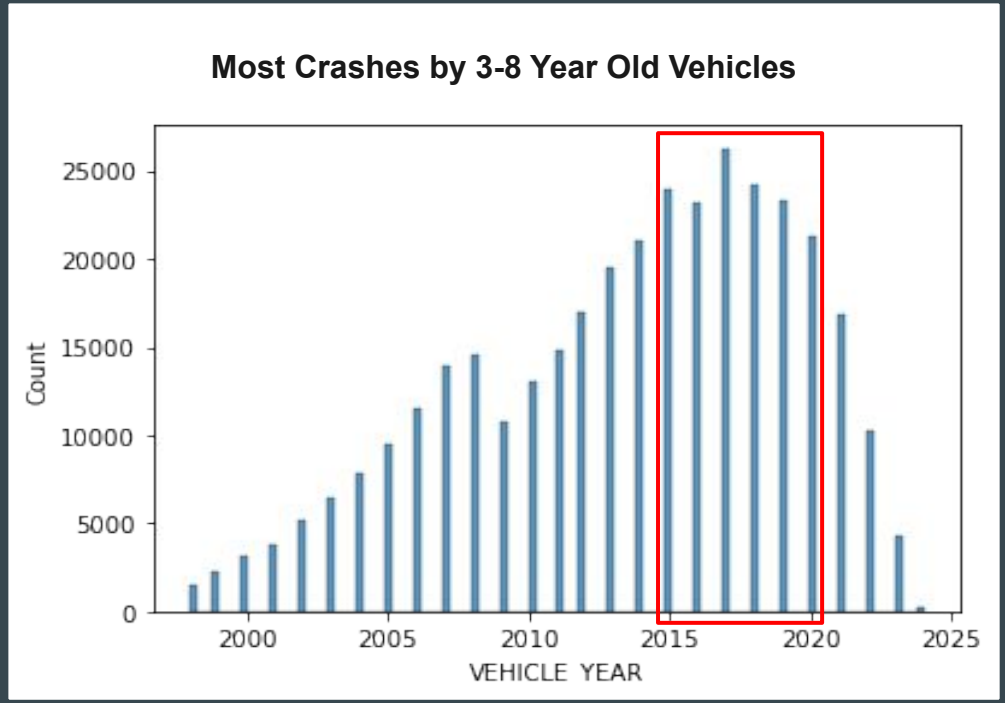
Analysis


Most Crashes by Ages 25-30
Count
10000
8000
6000
4000
2000
0
20
40
60
80
100
AGE

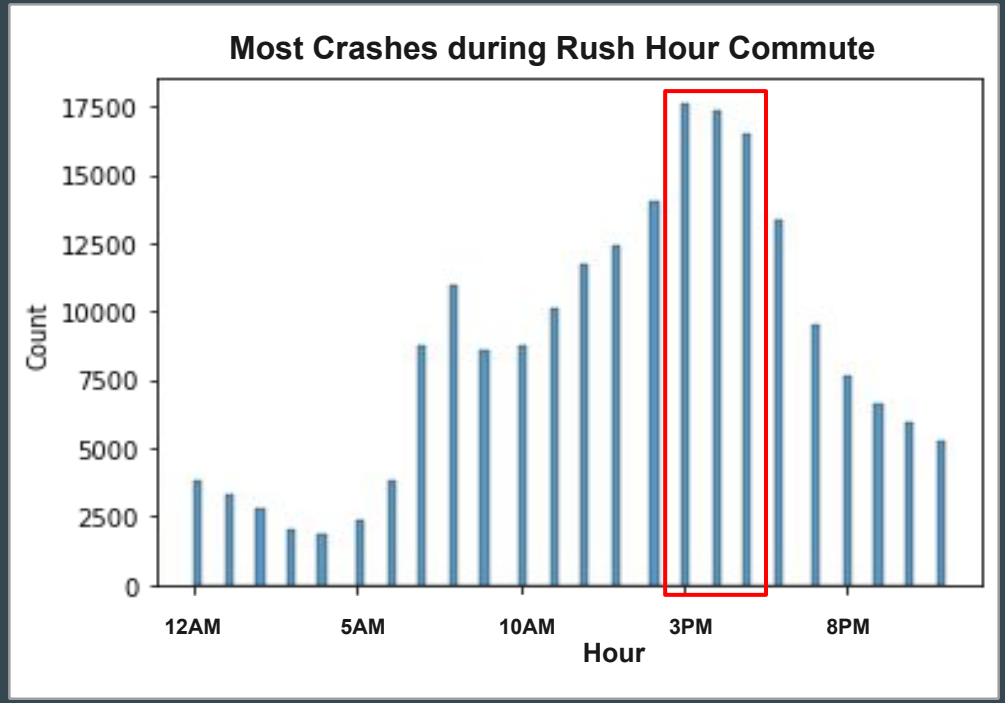
Analysis


55% of Crashes by Males
100000
80000
60000
40000
20000
0
Count
Male
Female
Unreported
Gender

Analysis


Most Crashes by 3-8 Year Old Vehicles
Count
25000
20000
15000
10000
5000
0
2000
2005
2010
2015
2020
2025
VEHICLE YEAR

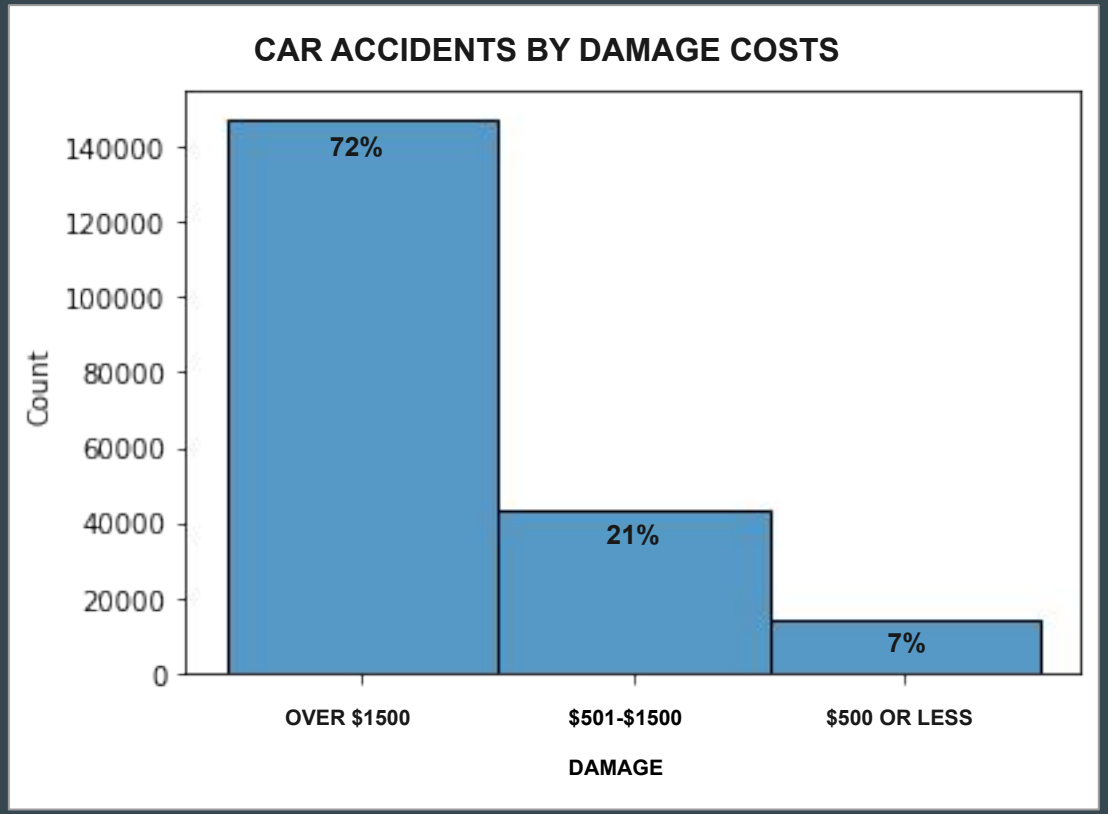
Analysis


Most Crashes during Rush Hour Commute
17500
15000
12500
10000
7500
5000
2500
0
Count
12AM
5AM
10AM
3PM
8PM
Hour

Analysis

Feature Engineering

- ❑ Target: Damage Over $1500
- ❑ Locals vs out-of-state
- ❑ Driving skills as a factor
- ❑ Time of day

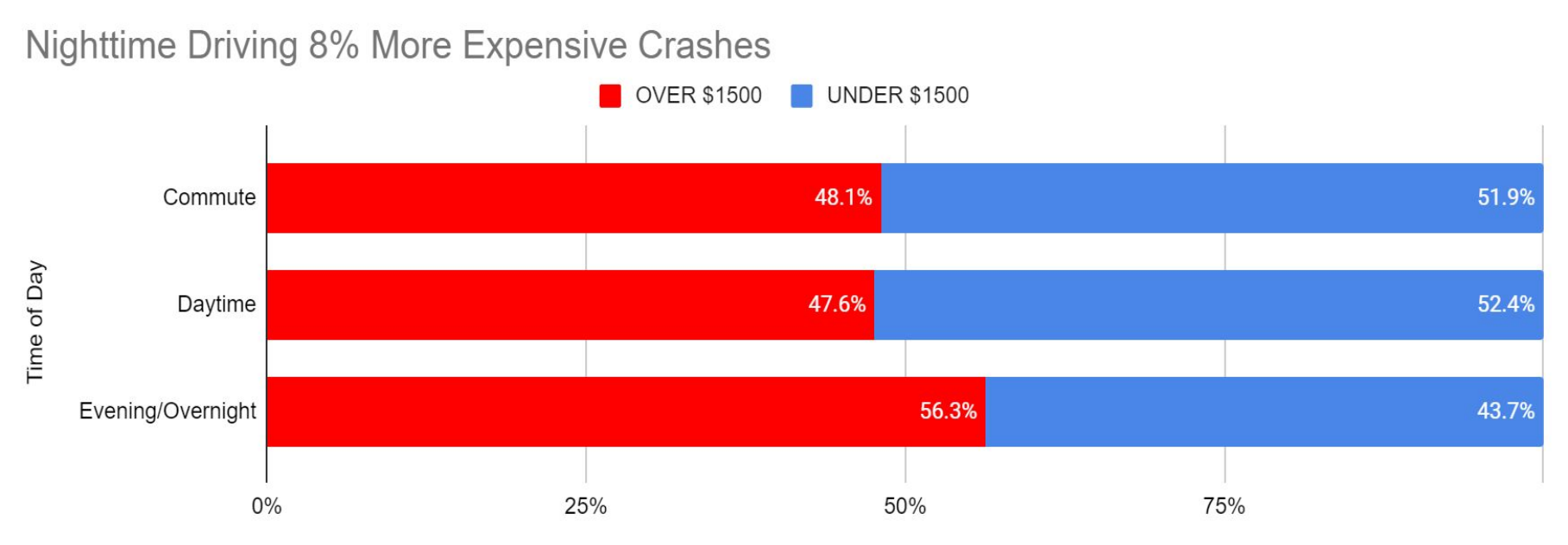


Analysis

Modeling

- ❑ Initial modeling indicated Time of Day
- ❑ Balanced Data
- ❑ Time of Day remained most influential

Analysis


Nighttime Driving 8% More Expensive Crashes
OVER $1500
UNDER $1500
Time of Day
Commute
Daytime
Evening/Overnight
48.1%
51.9%
47.6%
52.4%
56.3%
43.7%
0%
25%
50%
75%

Recommendations

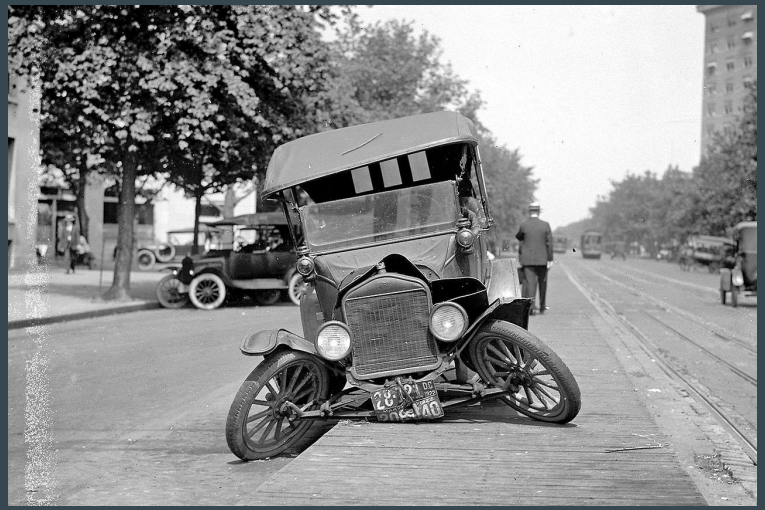
Recommendations

- Pay attention to the time of day.
- Offer customers an app to track driving data.
- Broaden business problem, focusing on driver age and gender, as well as vehicle age.

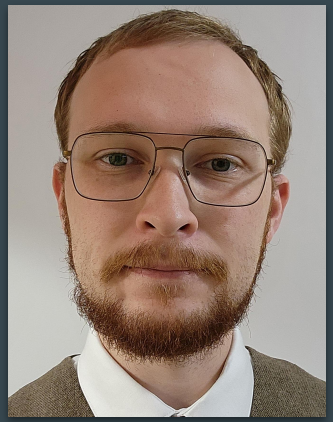
Future Insights

Future Insights

- Compare with total drivers on road
- Cross-reference actual repair costs to improve data on damage
- Expand the study to other locations, not only large cities



Questions?



John Baumgartner

Technical Lead

Github: jbaumgit

Email: jtbpilgrim@gmail.com

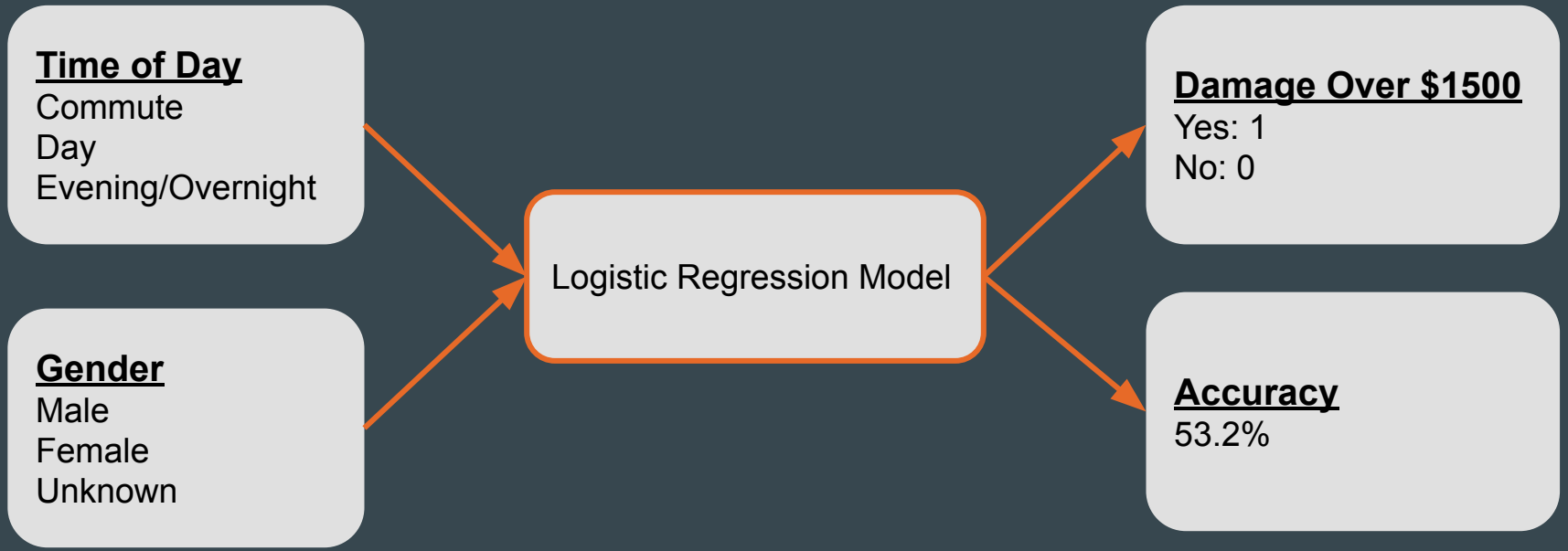
Dan Rosen

Github/Presentation Lead

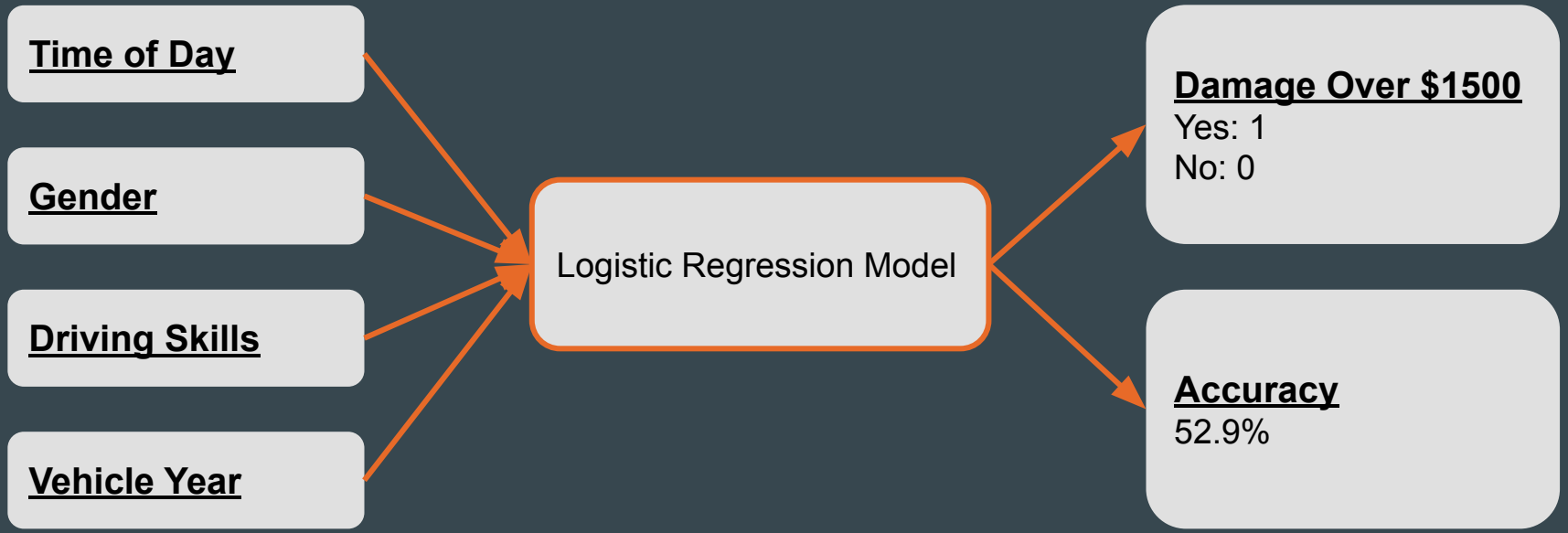
Github: dangrosen

Email: dan_rosen@outlook.com

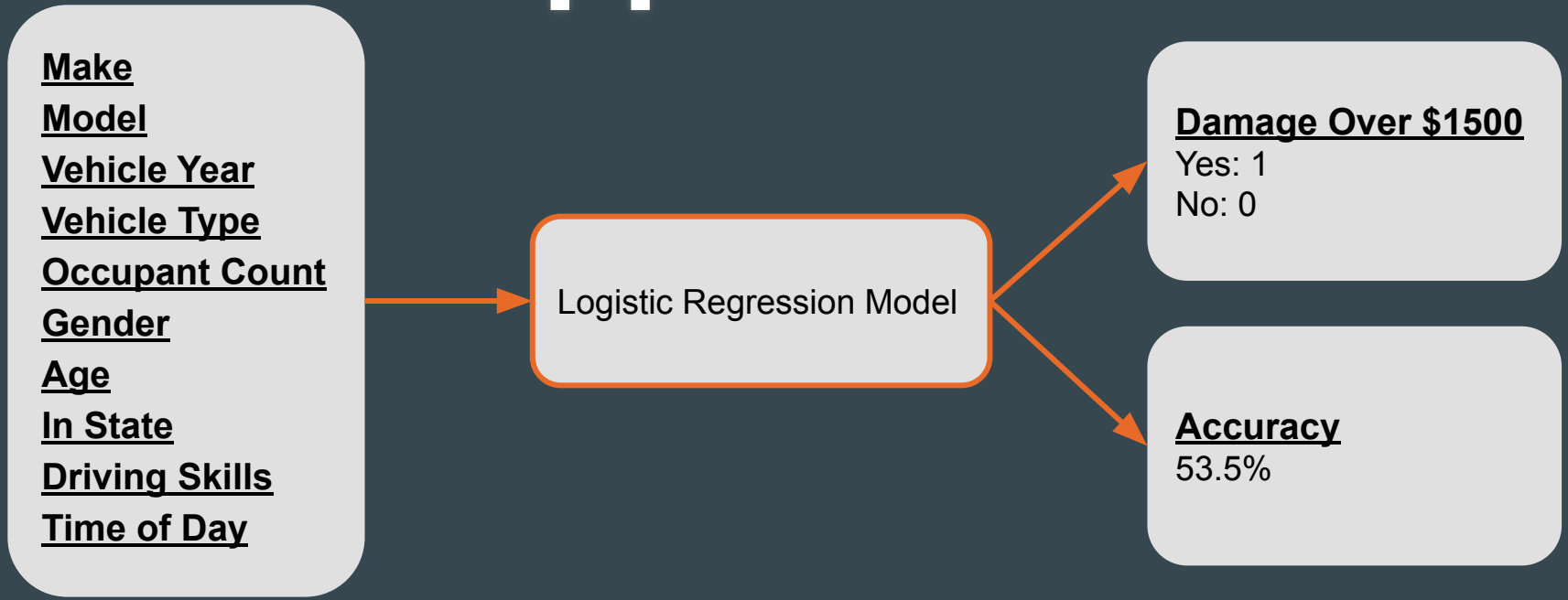
Appendix


Time of Day
Commute
Day
Evening/Overnight
Gender
Male
Female
Unknown
Logistic Regression Model
Damage Over $1500
Yes: 1
No: 0
Accuracy
53.2%

Appendix


Time of Day
Gender
Driving Skills
Vehicle Year
Logistic Regression Model
Damage Over $1500
Yes: 1
No: 0
Accuracy
52.9%

Appendix


Make
Model
Vehicle Year
Vehicle Type
Occupant Count
Gender
Age
In State
Driving Skills
Time of Day
Logistic Regression Model
Damage Over $1500
Yes: 1
No: 0
Accuracy
53.5%