Distribution of frequency of accidents in US

Hourly Distribution of Accidents

More severe accidents tend to affect longer road distance
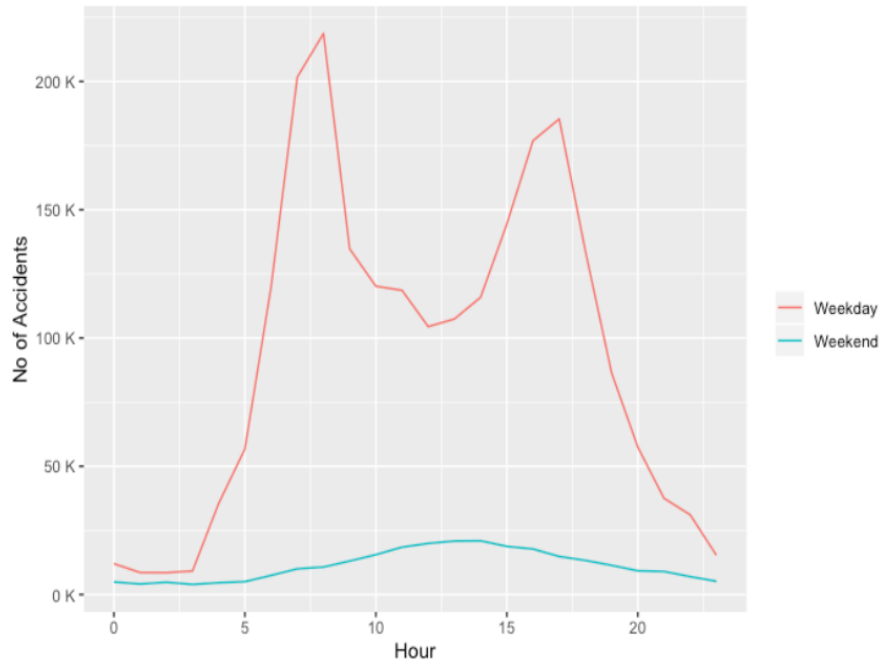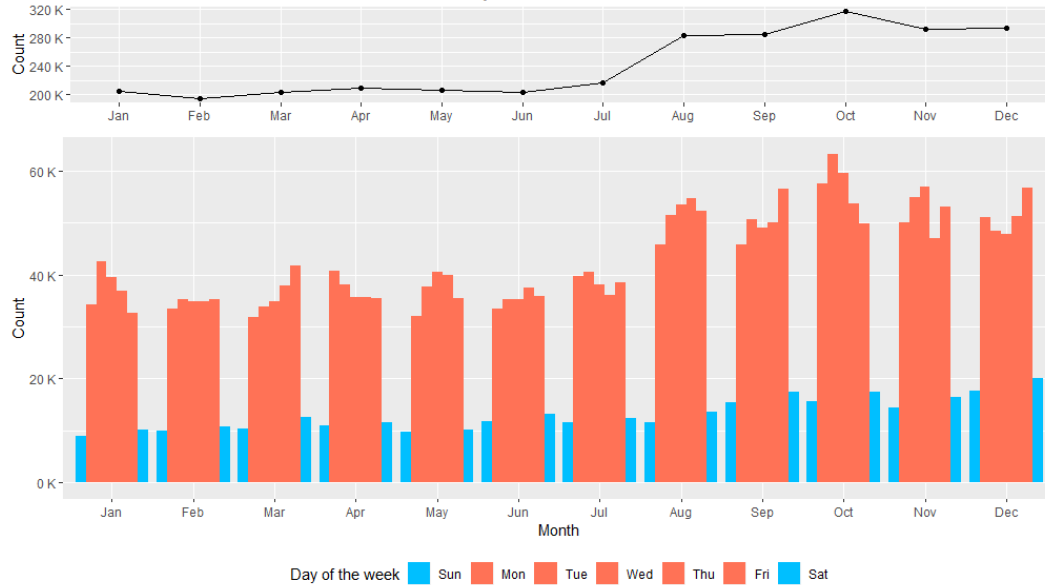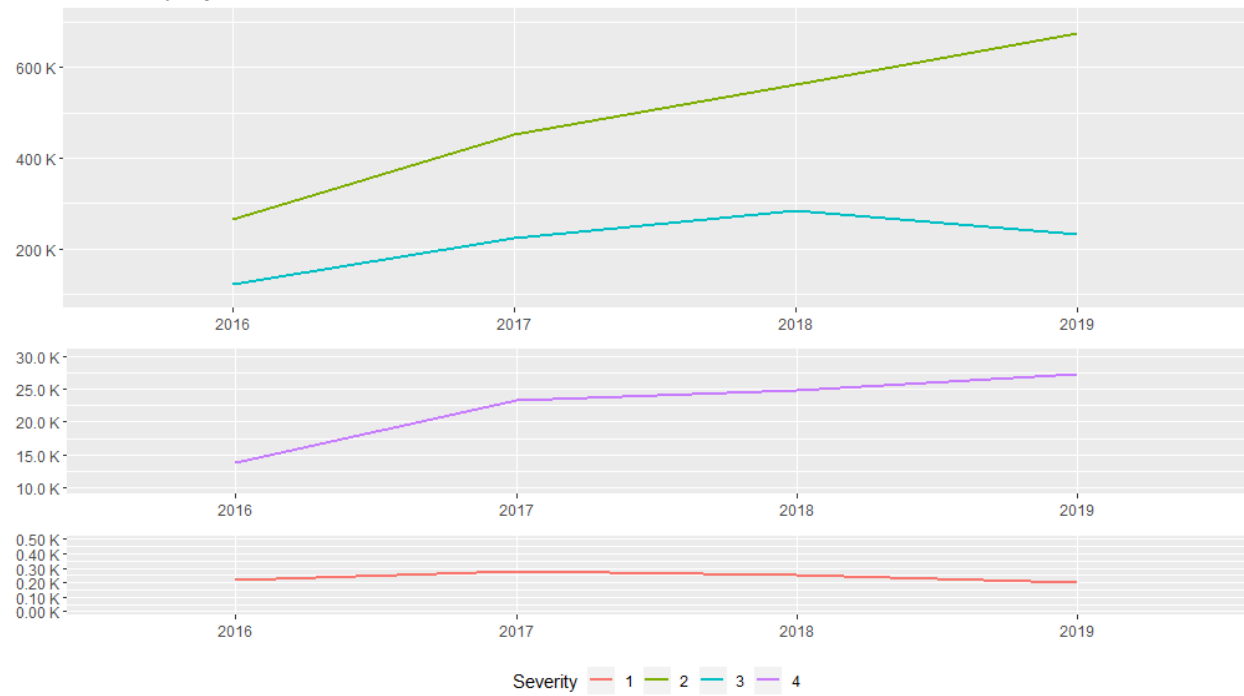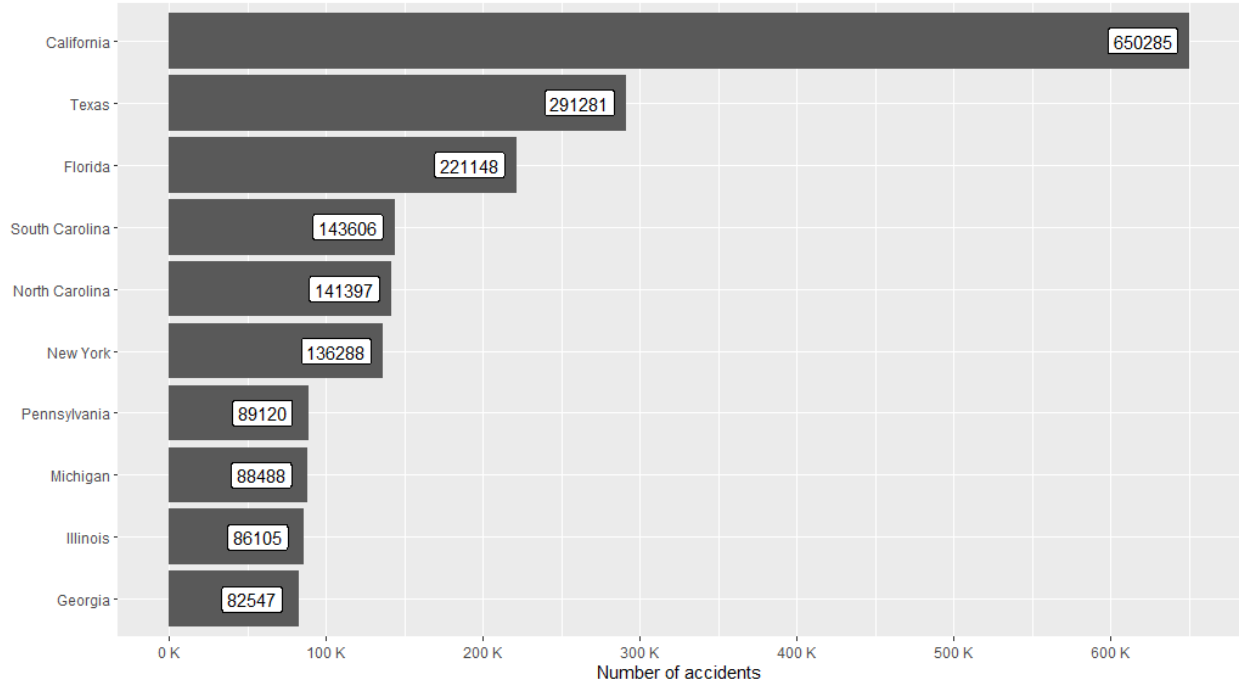
Hourly Distribution of Accidents

Pattern between accident counts and month & day of the week

Accidents per year

Severity — 1 — 2 — 3 — 4

Top 10 States with the most accidents

California	650285
Texas	291281
Florida	221148
South Carolina	143606
North Carolina	141397
New York	136288
Pennsylvania	89120
Michigan	88488
Illinois	86105
Georgia	82547

Number of accidents

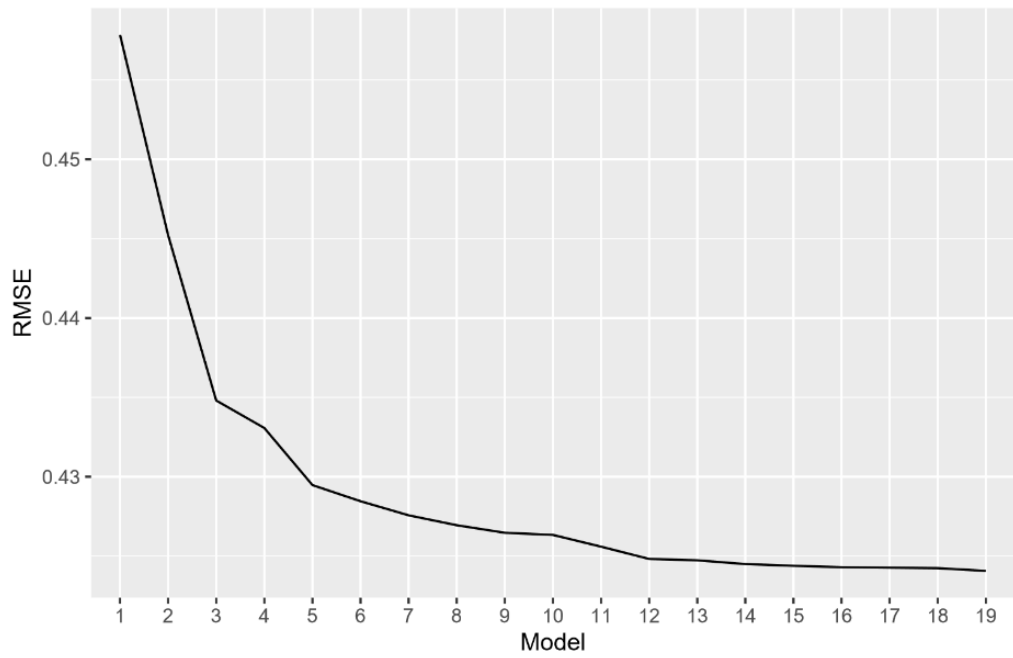## Duration of accident where Amenity shops exist

Accidents vs Duration(Minutes)
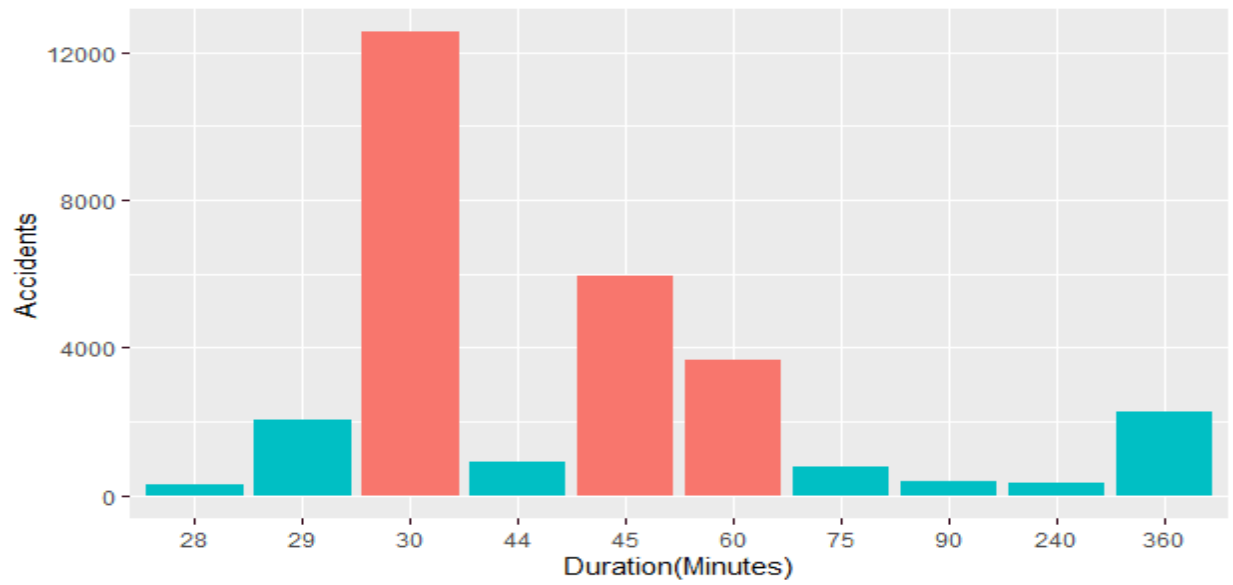
Duration of accident where Junction is present
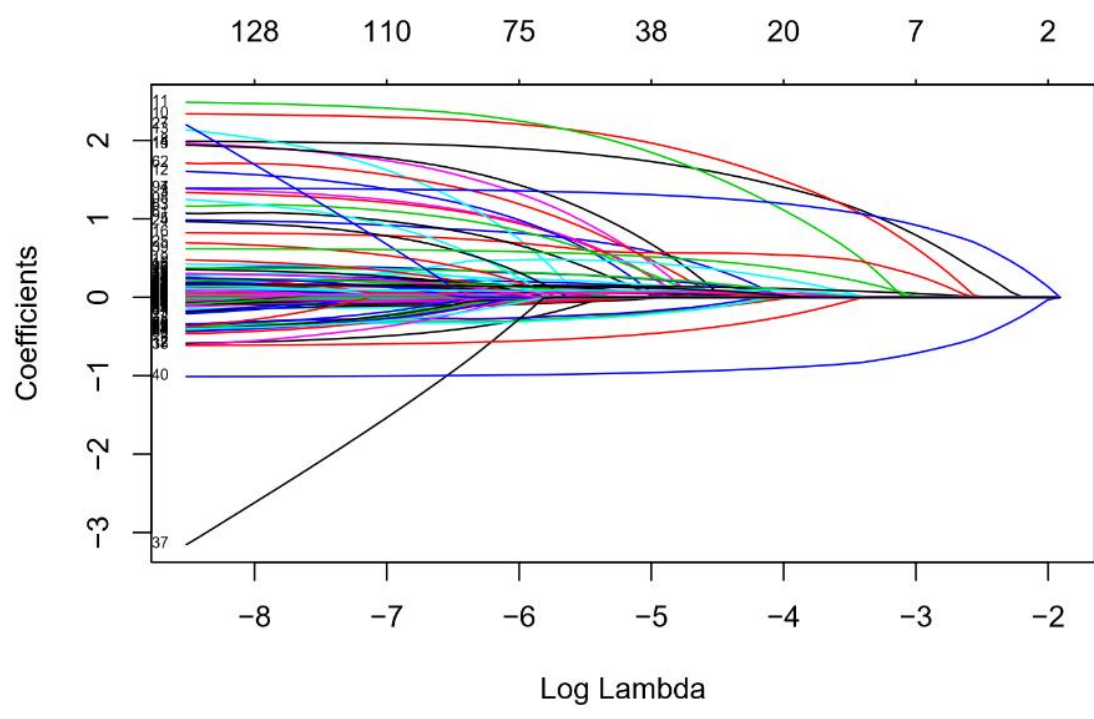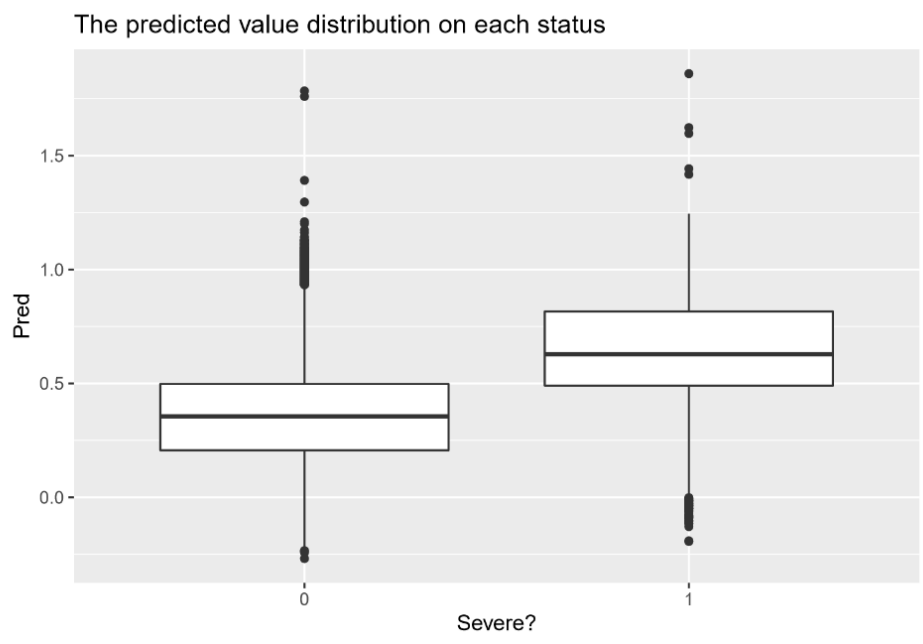
Accidents vs Duration(Minutes)
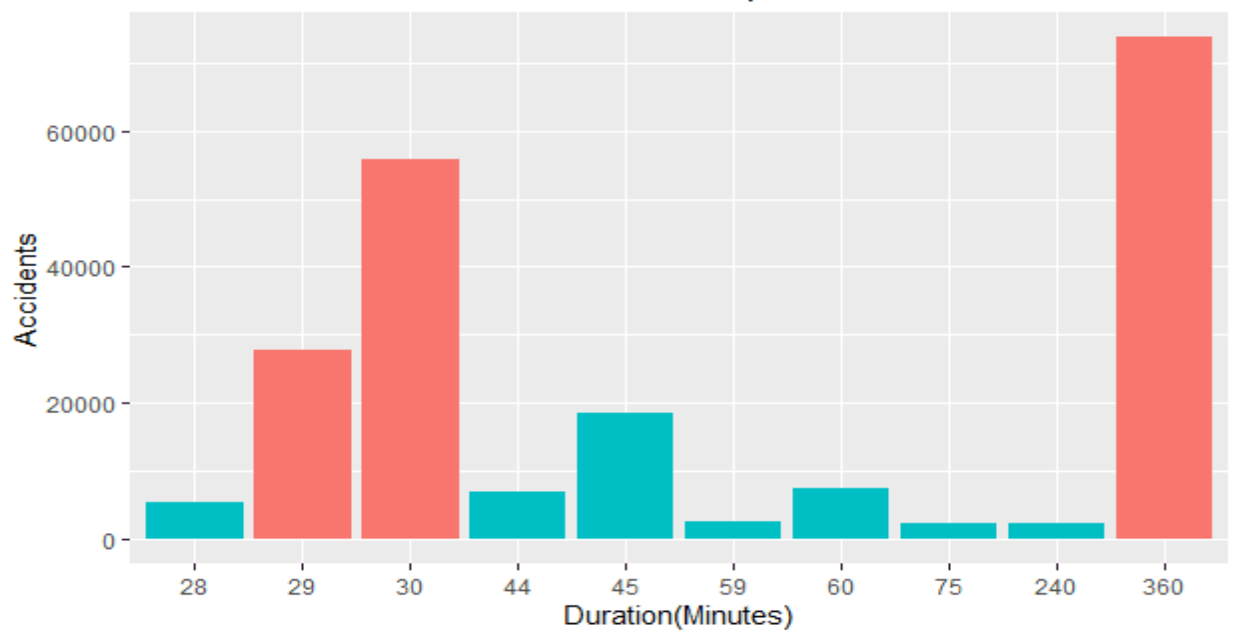
Duration of accident where Station/Railway is present

RMSE of each model
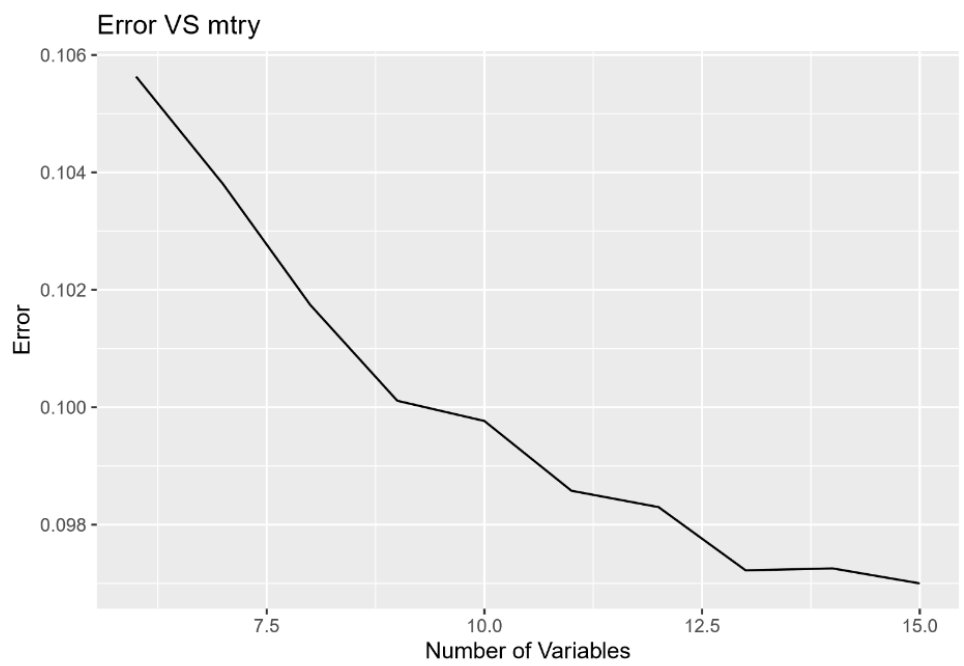
The predicted value distribution on each status

Error Rate

Error VS mtry
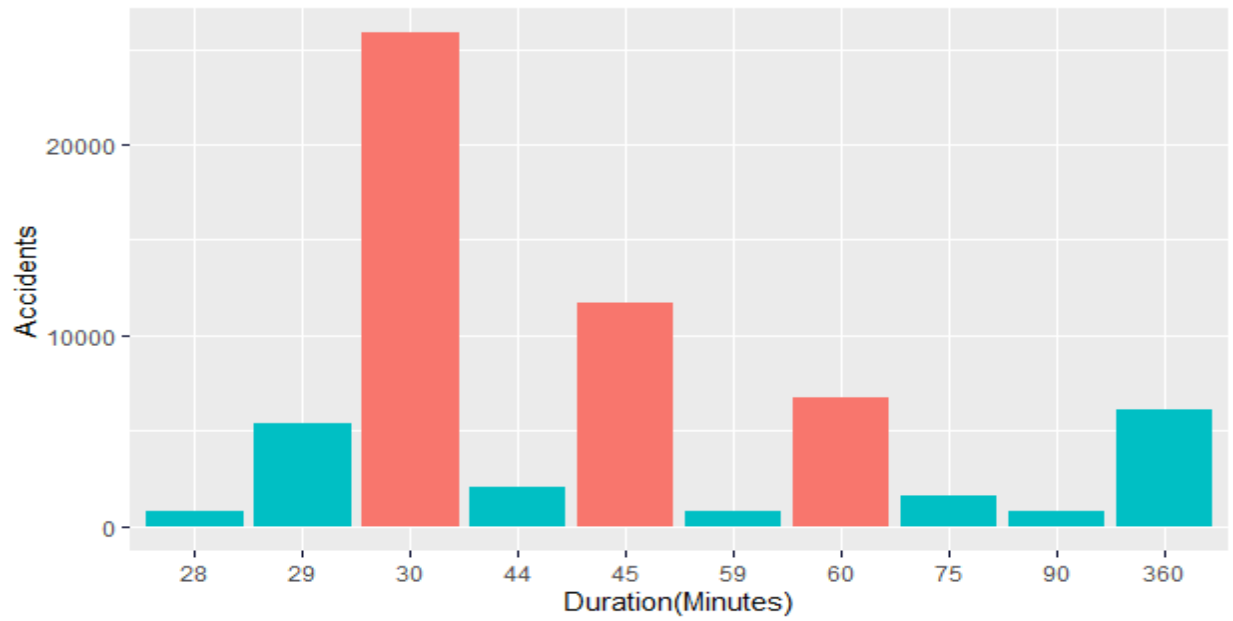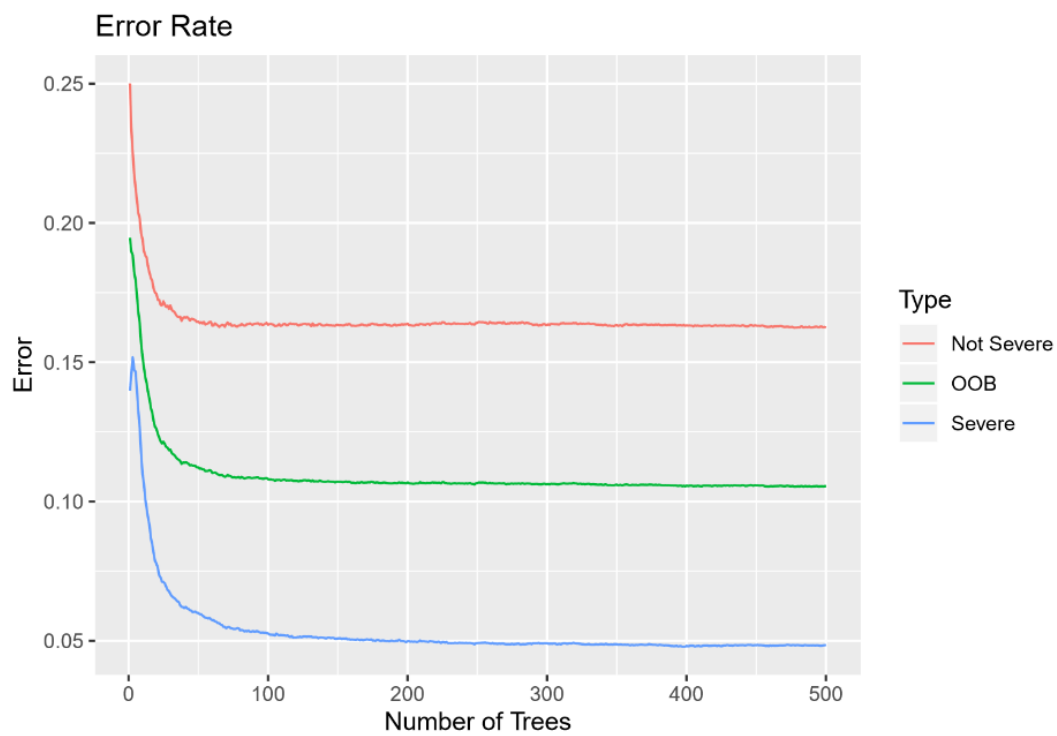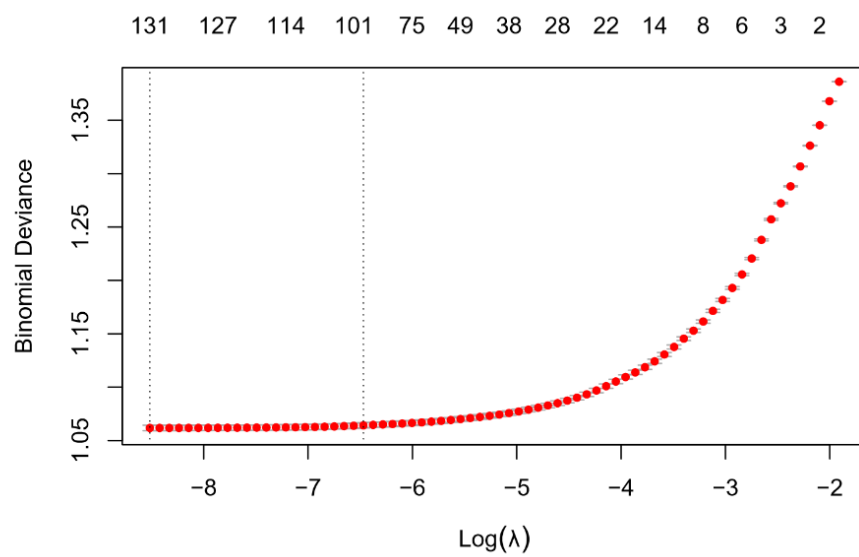
	logistic regression (cutoff is 0.5)	sparse logistic regression (cutoff is 0.5)	decision tree	random forest
accuracy	0.6723235	0.7086471	0.737066	0.8540312
sensitivity	0.6557434	0.6747925	0.7062519	0.8353701
specificity	0.7356831	0.7908939	0.8087898	0.8914362