



ACCIDENT SEVERITY PREDICTION

AMIT KUMAR MUKUL

24TH AUG, 2020

IMPORTANCE & TARGET

AIM: This project aims to predict the accident severity for given conditions.

TARGET:

- ❖ daily commuters who use their personal vehicle to travel in Seattle and
- ❖ Seattle Department of Transportation whose aim is to reach zero accident for an year

IMPORTANCE:

- ❖ Reducing the risk of meeting with an accident
- ❖ Decrease in Traffic Jams caused by Accident

DATA ACQUISITION & CLEANING

❖ Data was acquired from below site:

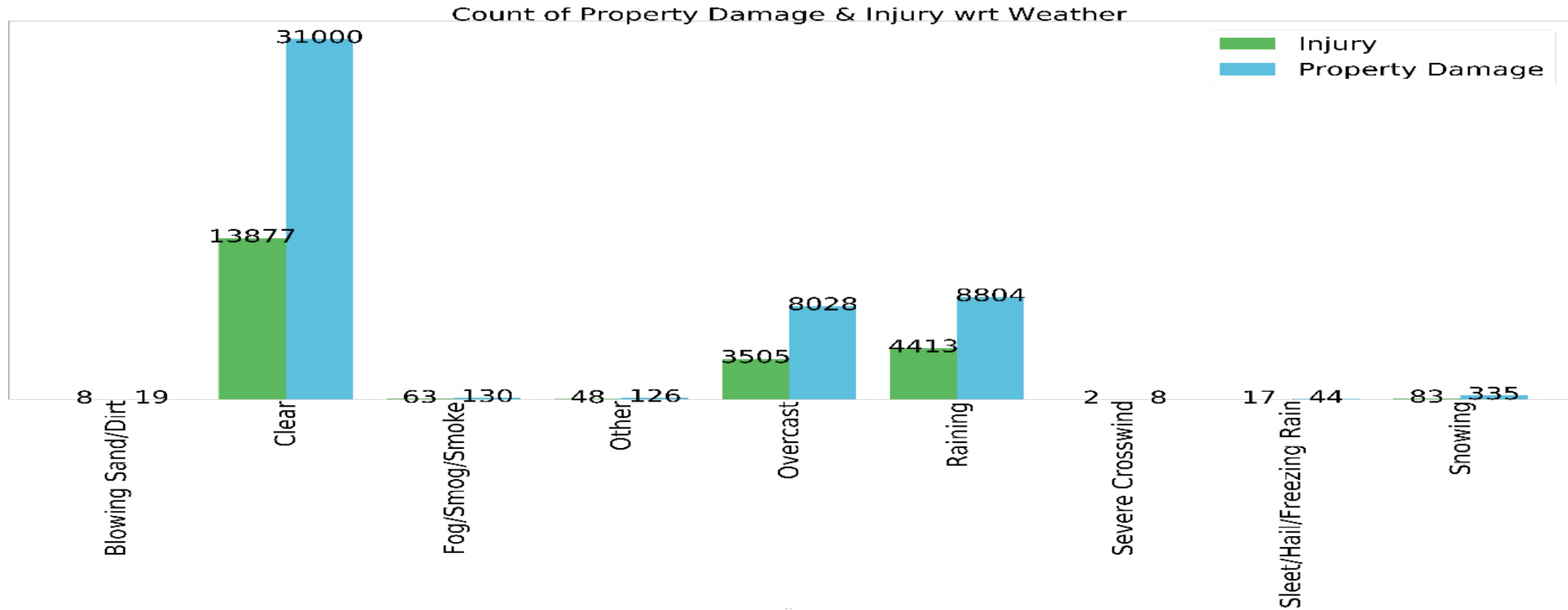
<https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv>

❖ Data consisted of 1946673 rows with 38 Attributes. Dataset was reduced to aprox. 70000 rows & 6 attributes. Attributes selection was done on the basis of exploratory analysis shown in subsequent slides.

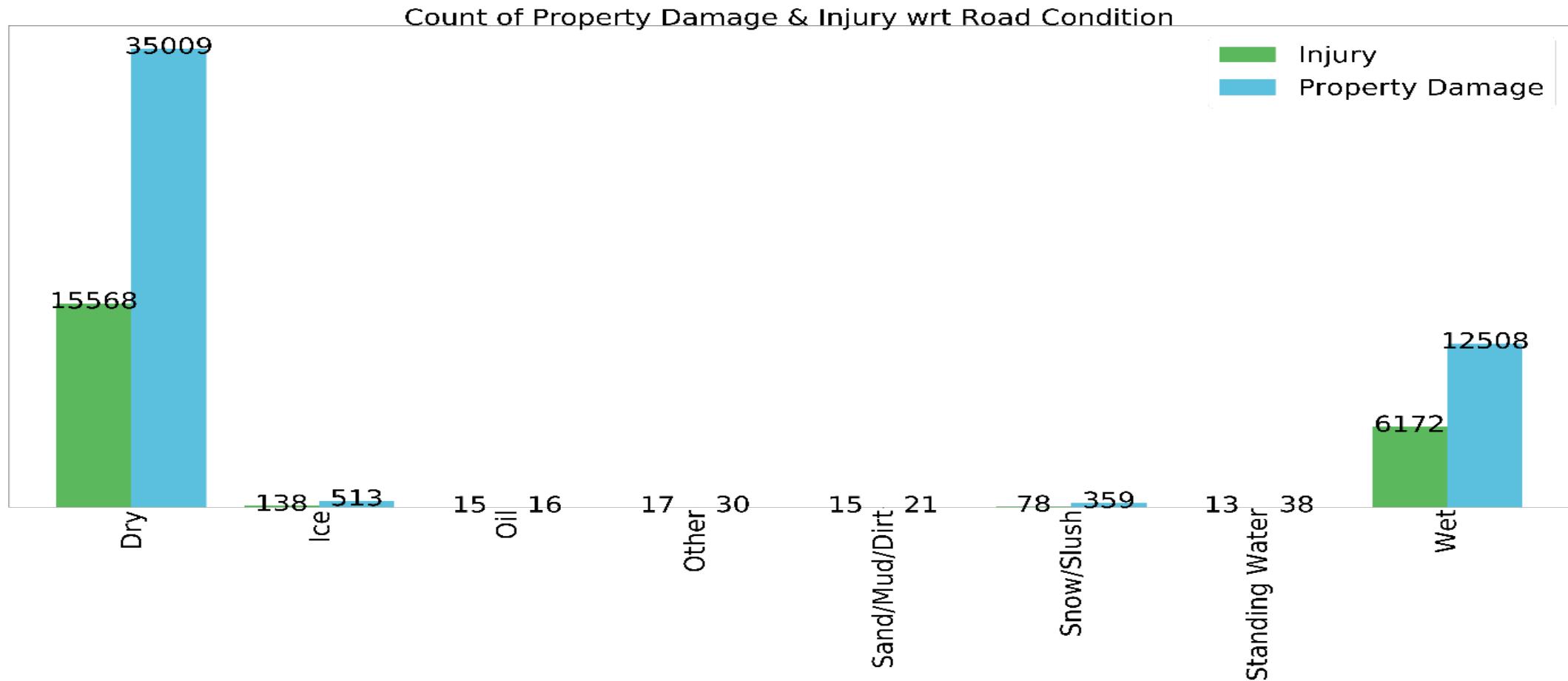
❖ Refined Dataset:

	SEVERITYCODE	INCDATE	UNDERINFL	WEATHER	ROADCOND	LIGHTCOND	Weeday_Name
0	2	2013-03-27 00:00:00+00:00	N	Overcast	Wet	Daylight	Wednesday
1	1	2006-12-20 00:00:00+00:00	0	Raining	Wet	Dark - Street Lights On	Wednesday
2	1	2004-11-18 00:00:00+00:00	0	Overcast	Dry	Daylight	Thursday
3	1	2013-03-29 00:00:00+00:00	N	Clear	Dry	Daylight	Friday
4	2	2004-01-28 00:00:00+00:00	0	Raining	Wet	Daylight	Wednesday

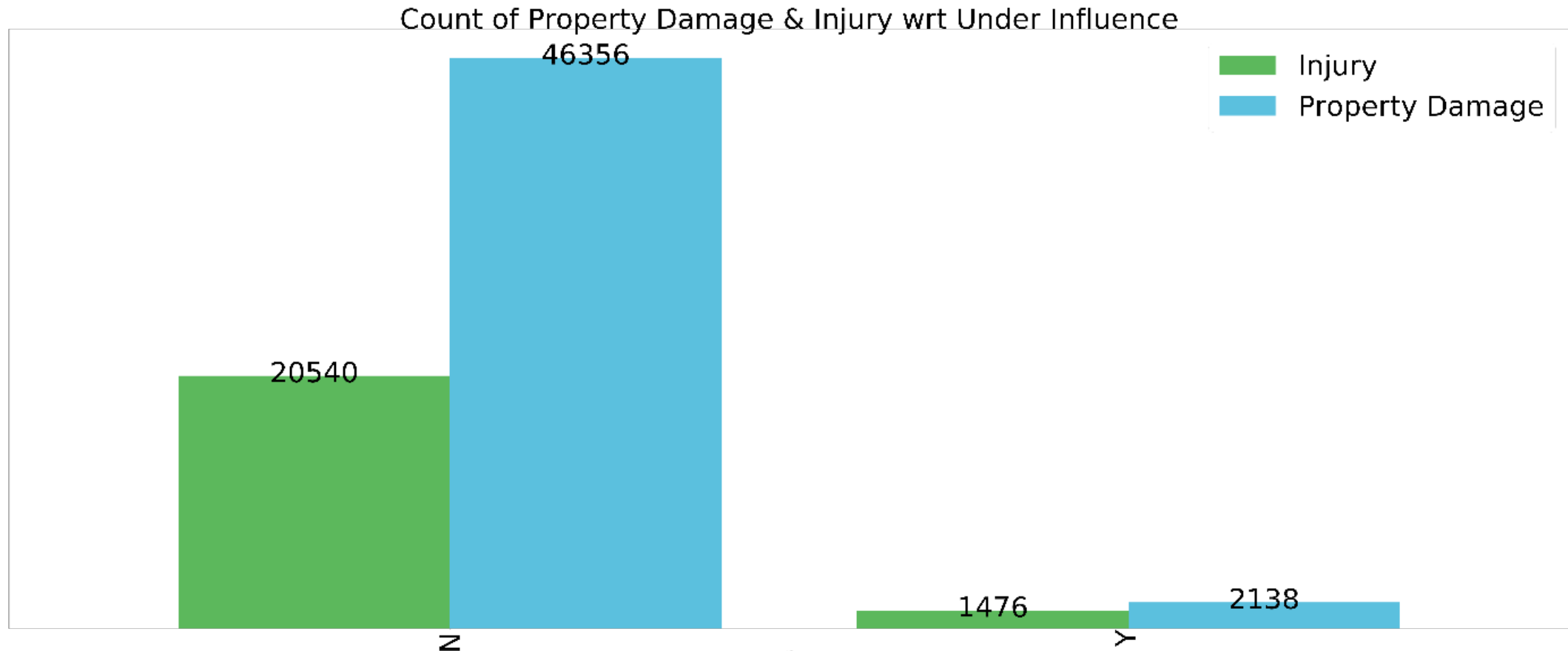
When it snows chance of getting Injured Reduces



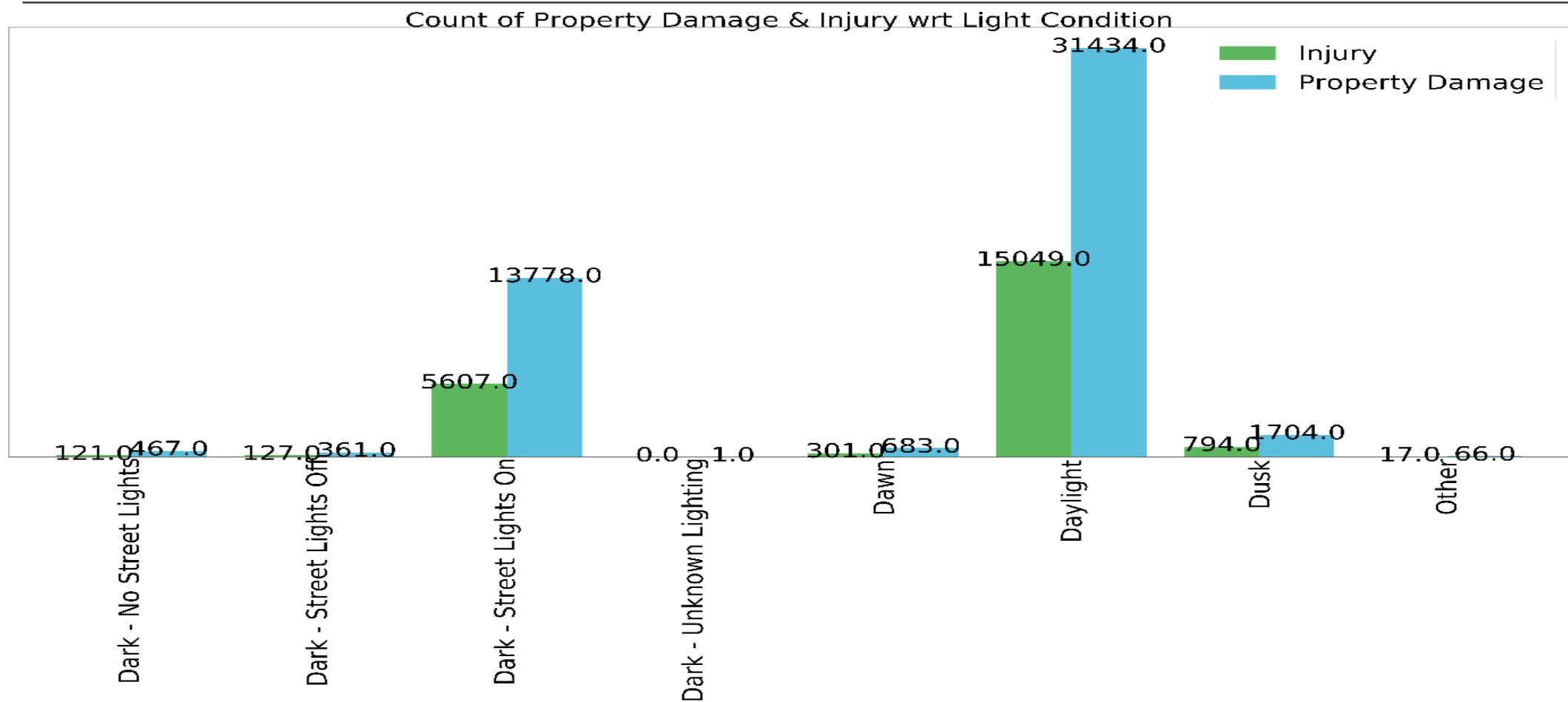
Higher chance of Injury on Oil & Other Category



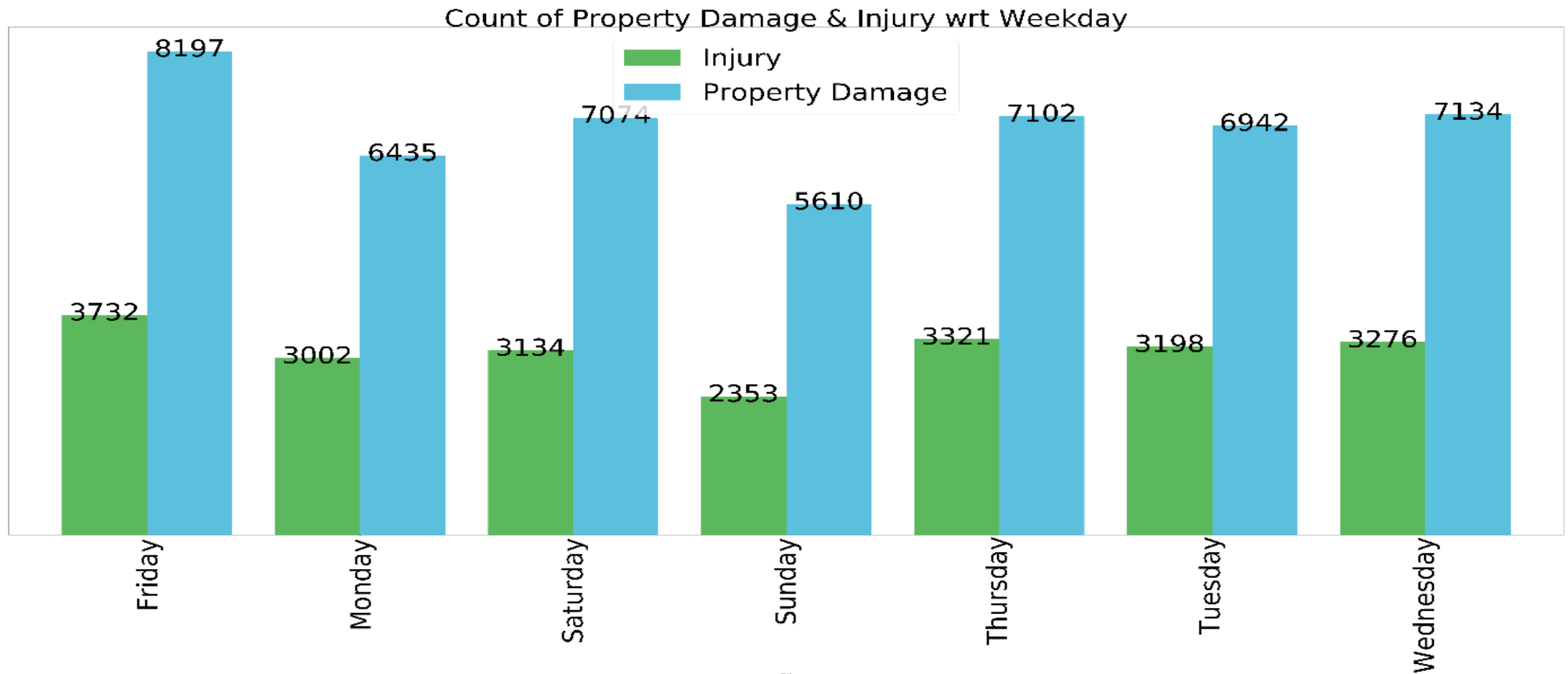
Influence of Drugs & Alcohol have Higher chance of Injury



Dark Roads have low chance of Injury



More Accident happens on Friday



Classification Model: Decision Tree

Evaluation Metrics	d = 1	d = 2	d = 3	d = 4	d = 5	d = 6	d = 7	d = 8	d = 9
Jaccard	0.686215	0.686215	0.686427	0.686286	0.686286	0.685789	0.686073	0.685435	0.685364
F1	0.558518	0.558518	0.559144	0.558945	0.559076	0.559226	0.559494	0.559183	0.559148

The Decision Tree model was finalised with Jaccard value of 0.686427.

Classification Model: Decision Tree

	Solvers				
Regularization Value	lbfgs	saga	liblinear	newton-cg	sag
0.1	0.619	0.618661	0.618660	0.618661	0.618661
0.01	0.618656	0.618656	0.618652	0.618656	0.618656
0.001	0.618627	0.618627	0.618953	0.618627	0.618627

Final Log Model was made from $C=0.001$ & liblinear

Classification Model: Support Vector Machine

Kernel Function	Accuracy Score
Sigmoid	0.558518008
Poly	0.558793187
rbf	0.558518008
linear	0.558642697

SVM model with Poly Kernel was finalized

Results & Conclusions

- ❖ All Models produced similar results with respect Jaccard Score and F1 Score.
- ❖ Relationship between Accident Severity and Variables affecting it was analysed.
- ❖ With data visualization it was really easy to see the relation of meeting with severe accident causing injury and dependent variables like Road condition, weather condition, light condition.
- ❖ This model can be run by SDOT officials or people who commutes in the city and they can easily identify if given condition can cause accident or not.

Thank You