

# **PREDICATING ACCIDENT SEVERITY**



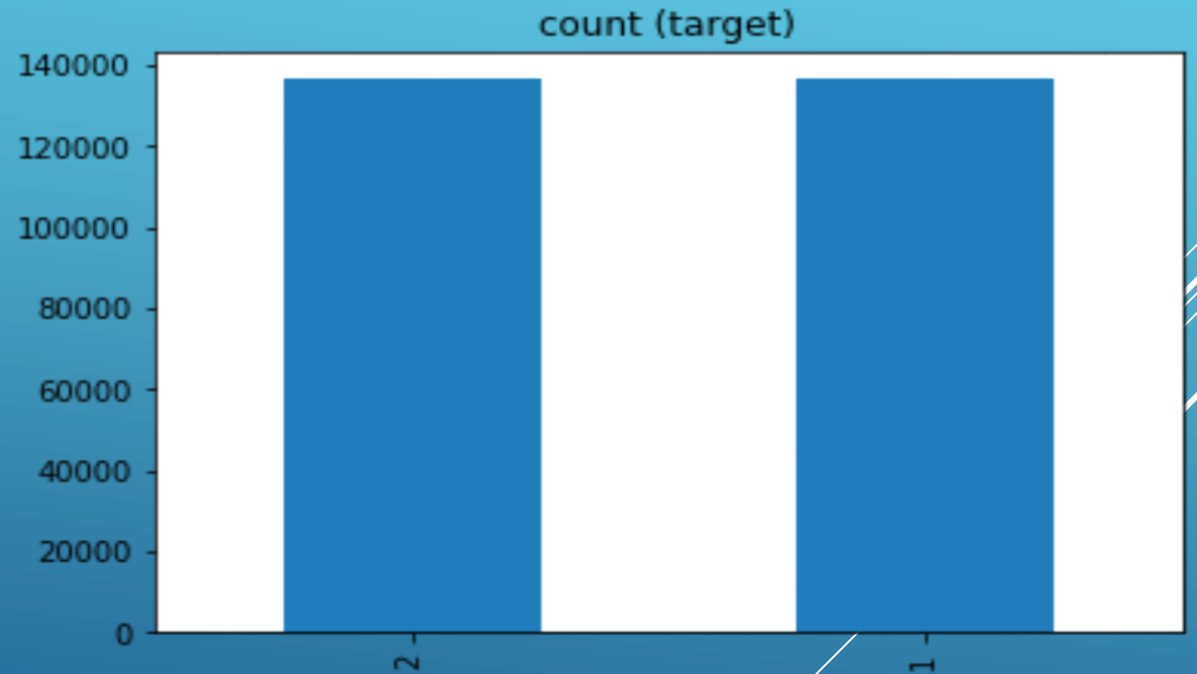
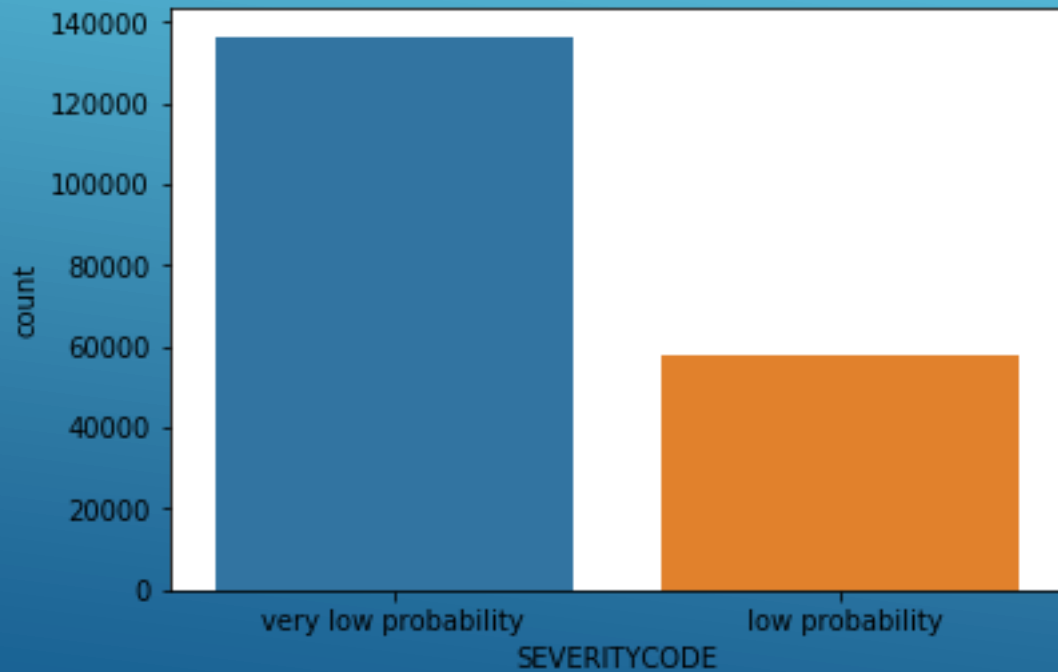
## **Introduction:**

- Road accidents are a major world economic and social problem as shown by the report of loss of lives and properties in many countries around the world.
- This study was designed to investigate the relative accident risk of different road condition, weather condition and road light condition and combinations of conditions.
- The main purpose for the project to provide advices for target audience such as police, rescue group, insurance companies and drivers themselves to make insightful decision it could help to reducing accident and accident severity.

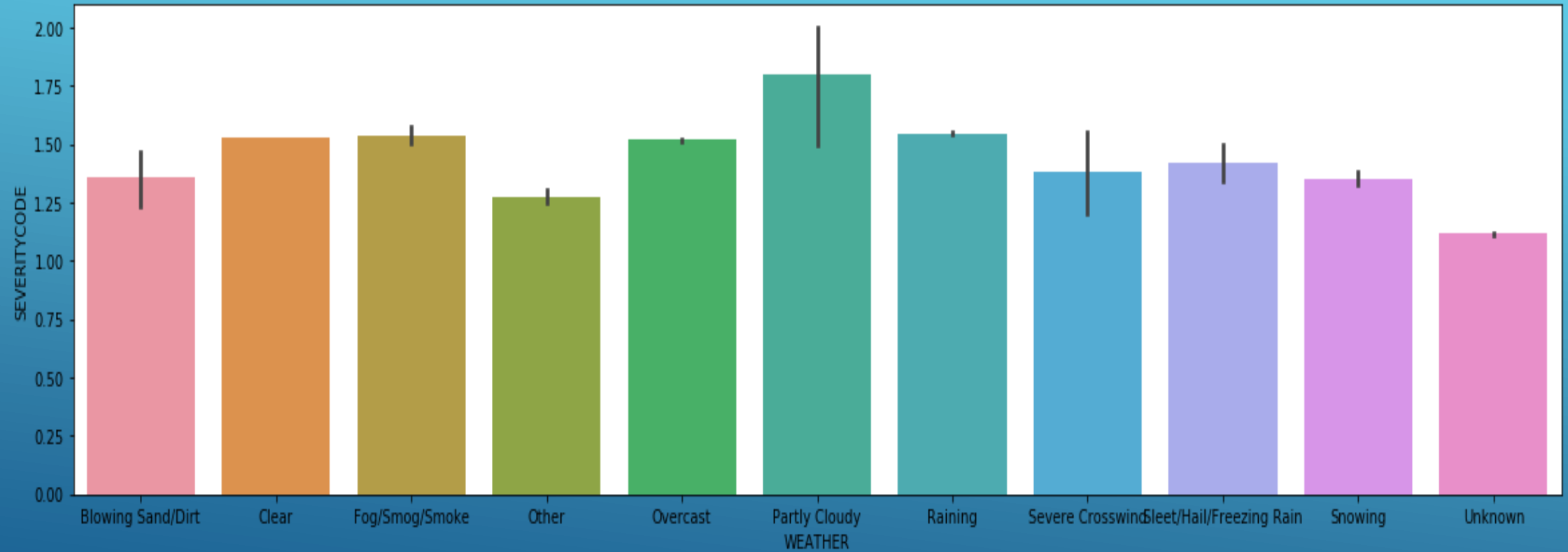
# Data acquisition and cleaning

- The Data provided by Seattle Police Department and Traffic Records department from 2004 to Present
- In total there are 37 attribute (independent variables) , The dependent variable (SEVERITYCODE) contains numbers of the correspond to different level of severity caused by accident from 0 to 4 .
- Some attribute have missing data.
- have numerical and categorical types of data need for preprocessing before any future processing.
- Unbalanced data so we will use some statistical method to balance it.

**Balancing the data by applying Random Over sampling by adding more copies to the minority class.**

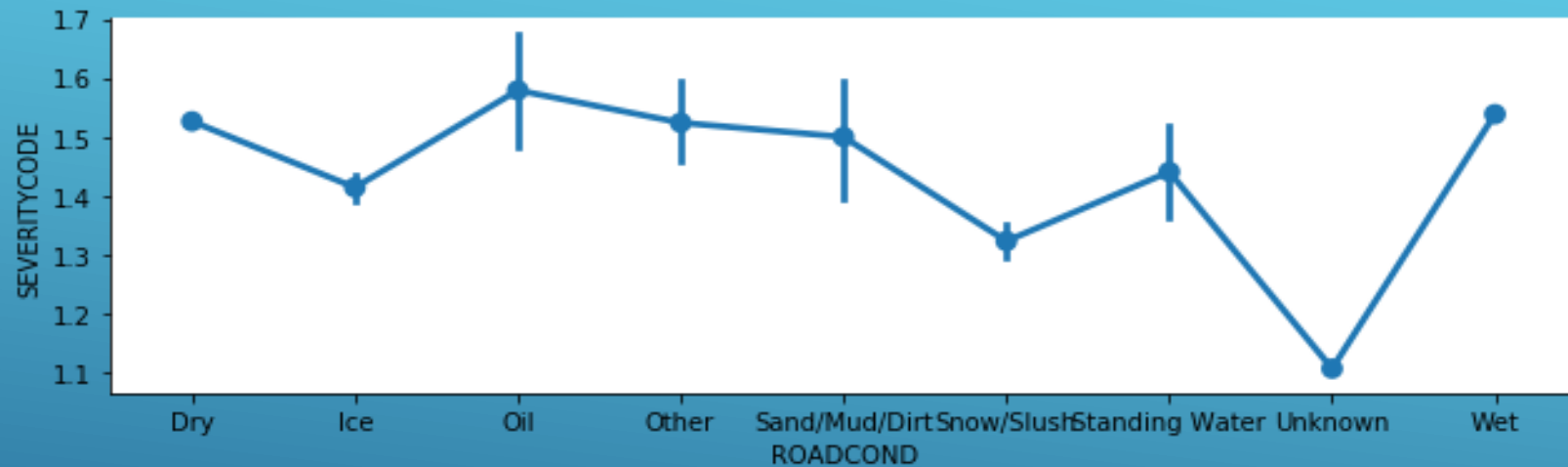


## D.1.Realtion between weather and accident severity



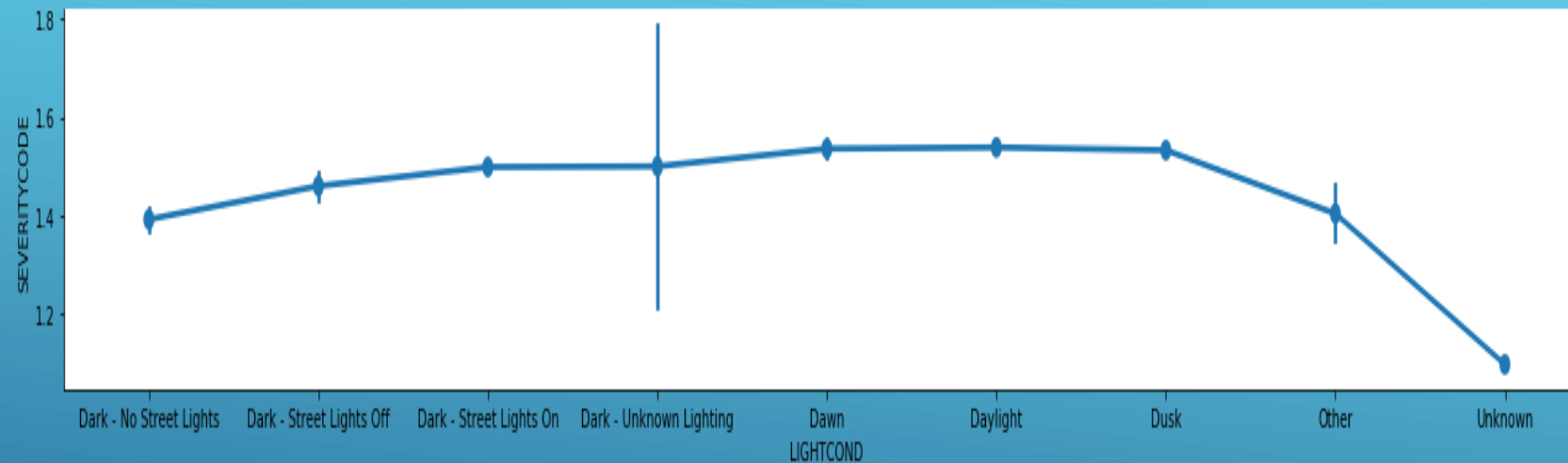
accident severity increase in partly cloudy weather compared with raining .

# Relation between road conditions and accident severity



The main factor which increase accident is oil and wet road.

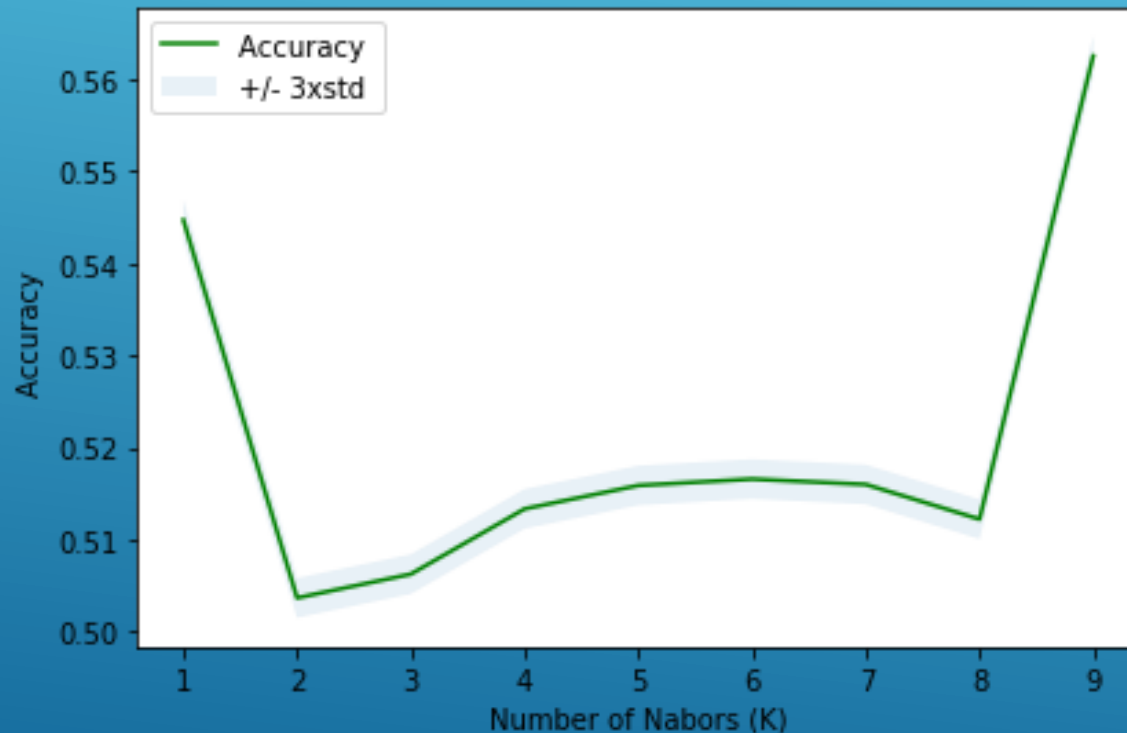
# Relation between Light condition and accident severity



Accident severity increase when it dusk and dark \_  
unknown lighting.

# Classification models

## K nearest neighbor(KNN)



The best accuracy was with 0.5607026413158955 with k= 9

Train set Accuracy: 0.561971095724805

Test set Accuracy: 0.5625343444334542

jaccard\_similarity\_score: 0.5625343444334542

f1\_score:0.5032204390014452



# Classification models

Performance of classification models.

	Algorithm	Jaccard	F1-score	Logistic Regression
0	KNN	0.550885	0.527015	Nan
1	SVM	0.564220	0.543300	Nan
2	Logistic Regression	0.529912	0.516090	0.684082

## **Conclusion**

- **Built useful models to predict accident severity .**
- **Accuracy of the models has room for improvement.**
- **Ideas include:**
  - **Road condition**
  - **Weather condition**
  - **Light condition**