CAPSTONE PROJECT
CAR ACCIDENT SEVERITY

Applied Data Science

Introduction

Traffic accidents are bad events to experience.

Frequent accidents cause many losses for many people. In addition to claiming lives, accidents also cause material losses because these events can result in traffic jams or damage to public infrastructure that have been the target of collisions so that national development is hampered.

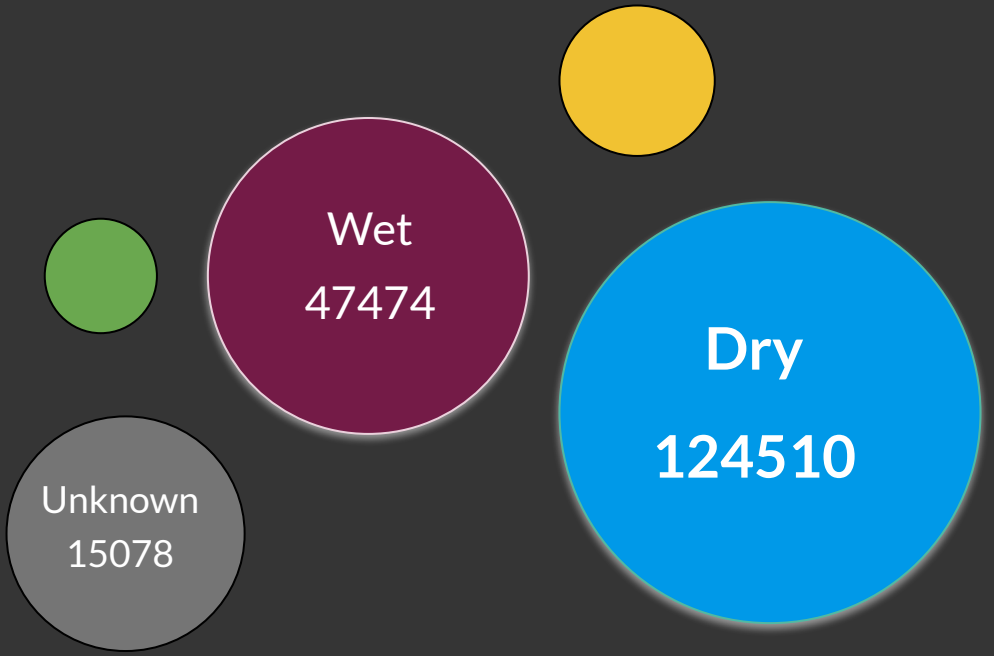
The Seattle government, through the Seattle Department of Transportation, has released data on accidents that have occurred in Seattle since 2004 and it is always updated every week. The uploading of the data shows how concerned the Seattle government is about this topic.

1. Intro

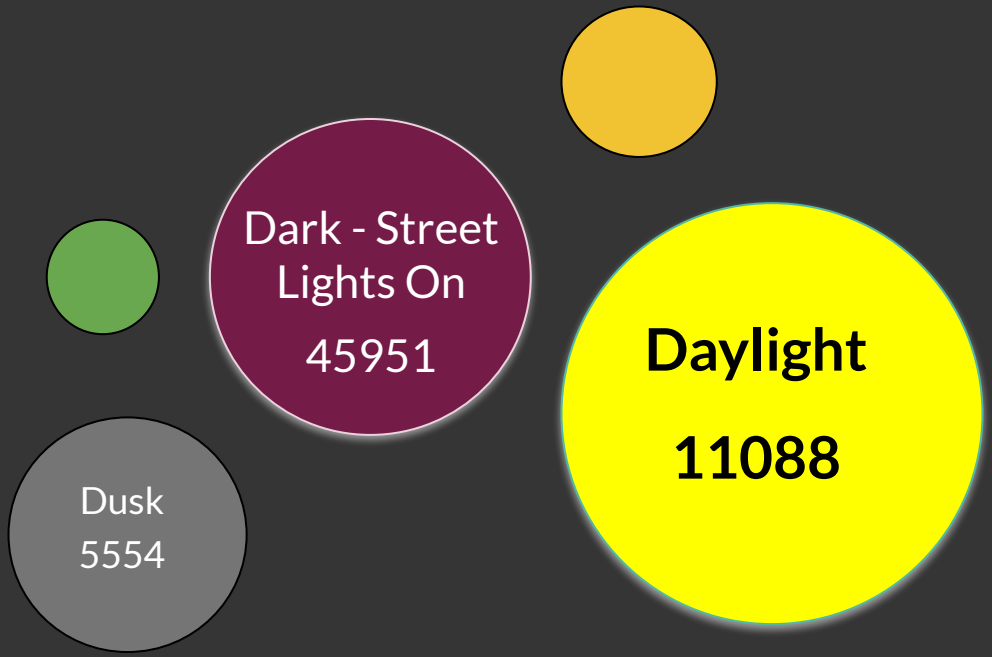
Reporting from **Colburn Law**, it is known that almost every day, there are collisions in Seattle. In 2016, collisions severity occurred as follows: **58%** properly damage only, **27%** injury collisions, and **1%** serious injury collisions. It is also known that the "common" type of collisions occurred were car collisions, with a total of 3,644 incidents in 2016.

Is it true car accident most occur when the sky is clear?

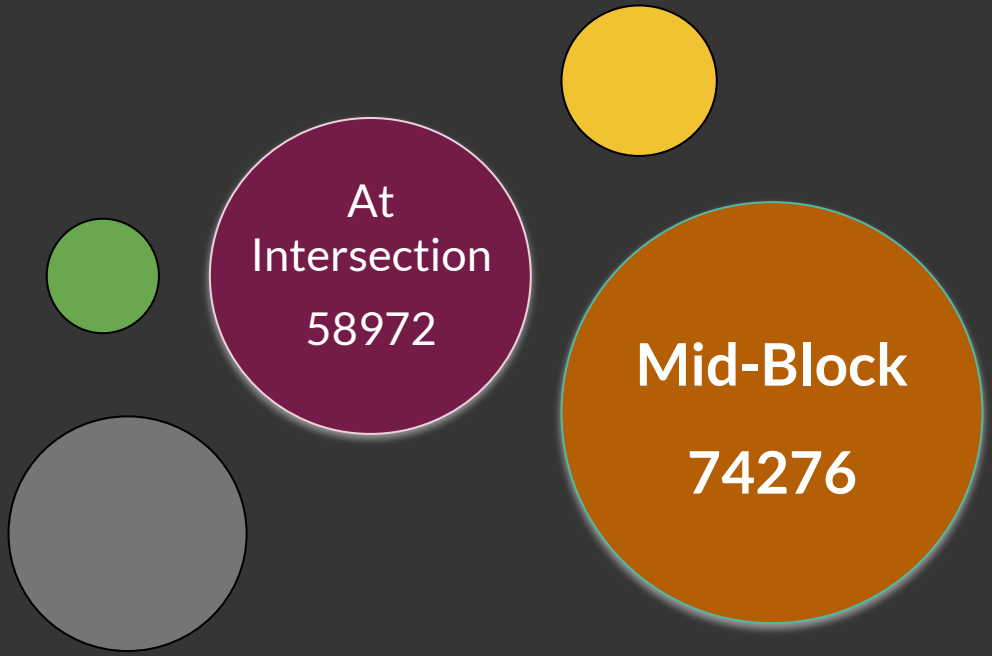
How about the road condition?



How about the light condition?

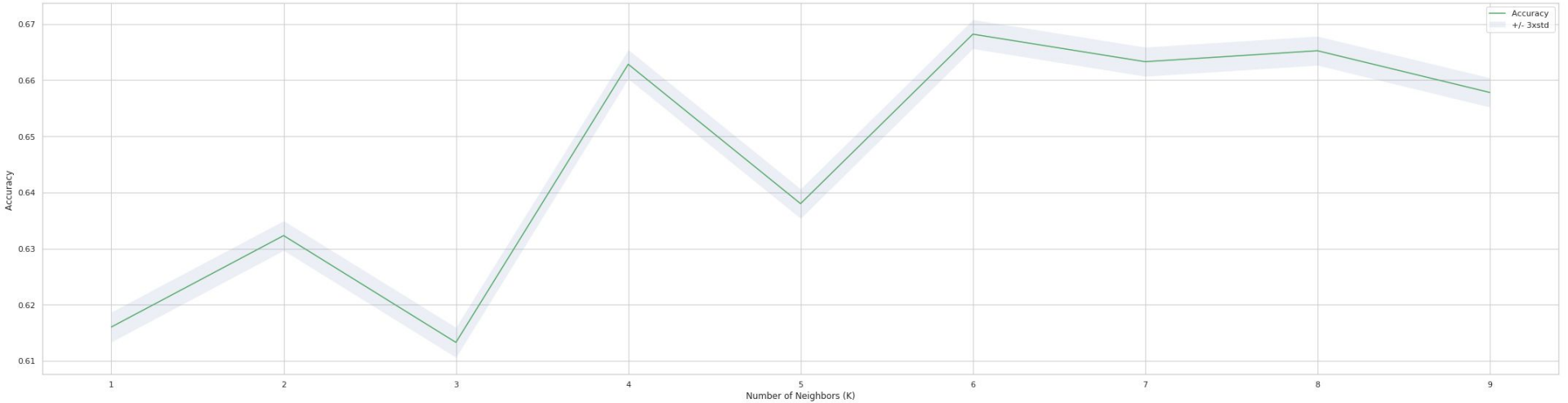


How about the junction type?



Modelling data using

KNN (K-Nearest Neighbor)



RESULT

- Accidents occur most frequently during the day with dry road conditions.
- This can happen because in both conditions it is a common time for people to have activities outside the home, such as school, work, and so on.
- The model can carry out car accident severity prediction (at k = 8) with a score of $R^2 = 0.66$
- Good enough number so that it can be said that this model can make predictions with the desired variables.

With the existence of data that has been published by the government and other authorities, it can make it easier for data scientists around the world to carry out.

in the next, regression techniques can be applied to make the car accident severity prediction program more real.

Conclusion

machine learning can be applied to real issues in the social environment,

including in terms of predictions.

KNN shows that this modeling has a fairly good level of accuracy.