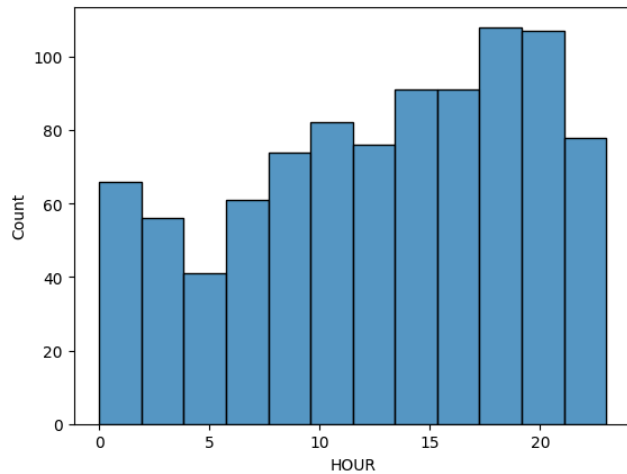


Explore Data Results

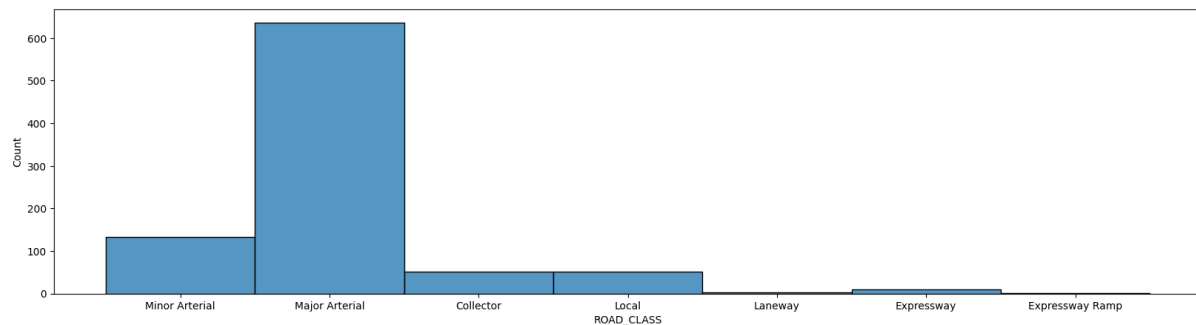
The size of the dataset was pretty good for the sake of any modeling to be done which we discuss further. The dataset contains about 931 instances and we closely looked through the entire data to identify which out of the 57 attributes could be useful for the purpose of this project.



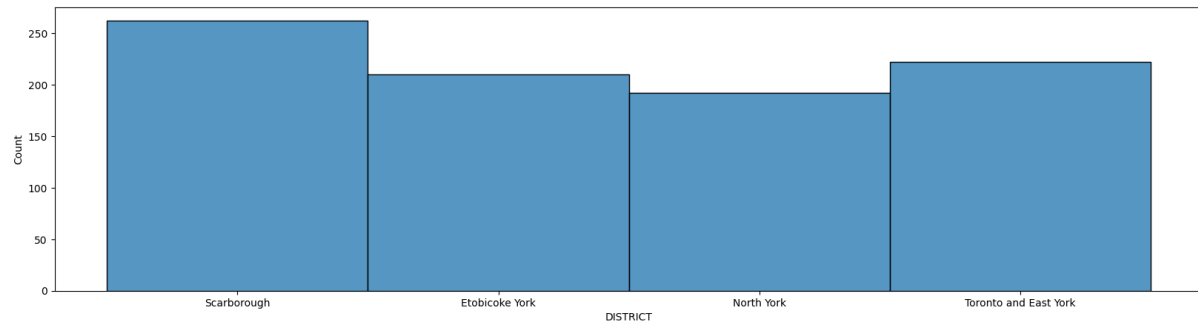
Accidents were the most during late hours where they were about 100 or more accidents

Further explorations were carried out to understand how each or any attribute could be deterministic to our problem goal.

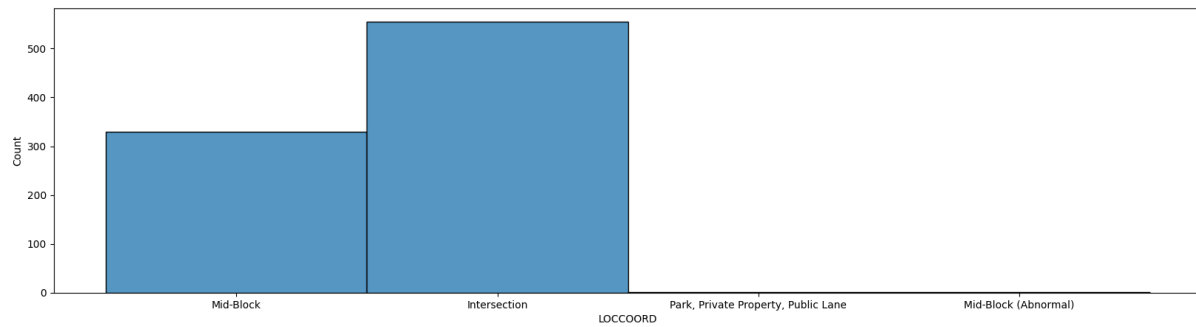
We explored various attributes of each and every attribute starting with which roadclass was very prominent for accidents and identified a few.



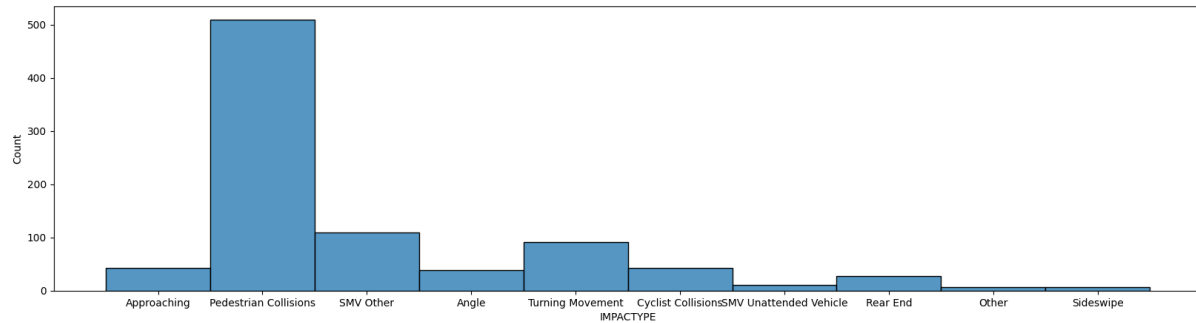
Majority of actions happened in Major Arterial road class



Scarborough has the highest number of accidents

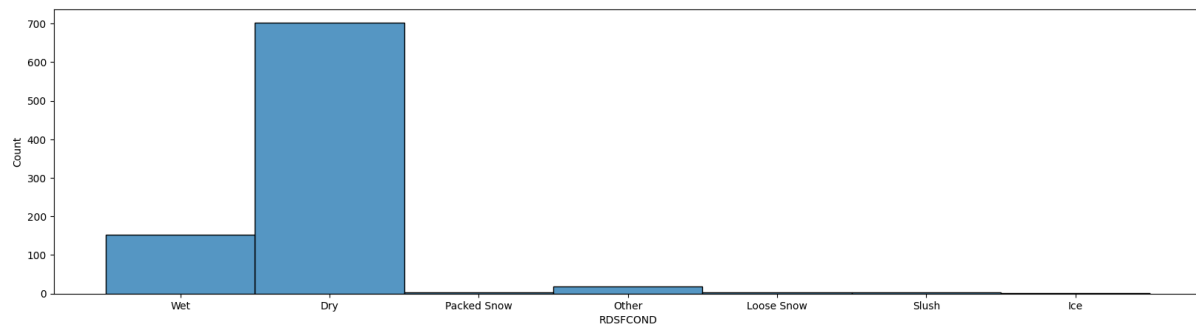


Intersection was where most of the accidents happened.



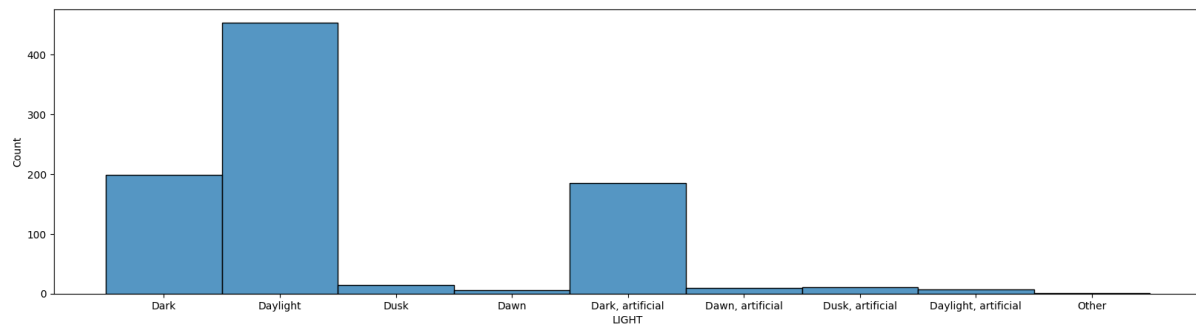
Impact type of Pedestrian collisions were the highest among the count.

We expected that most of the accidents would be in snowy weather but most of the accidents happened during dry weather.



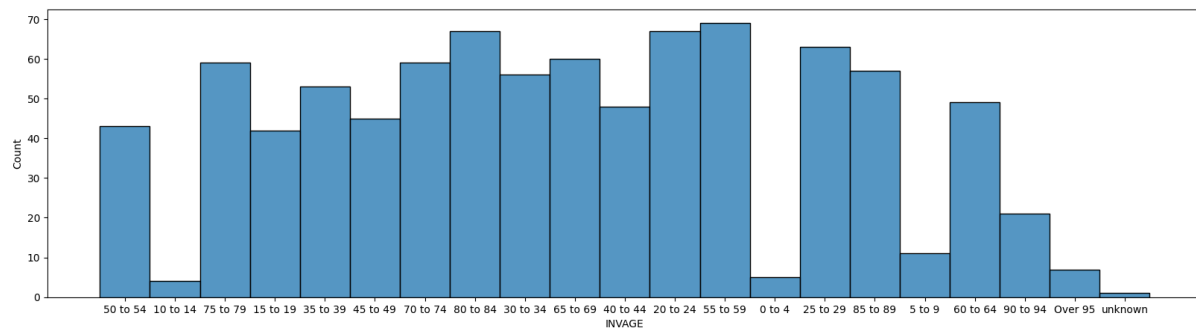
Dry weather with the highest number of accidents.

We tried to identify the lighting conditions during most of the accidents and Daylight was when most accidents happened.



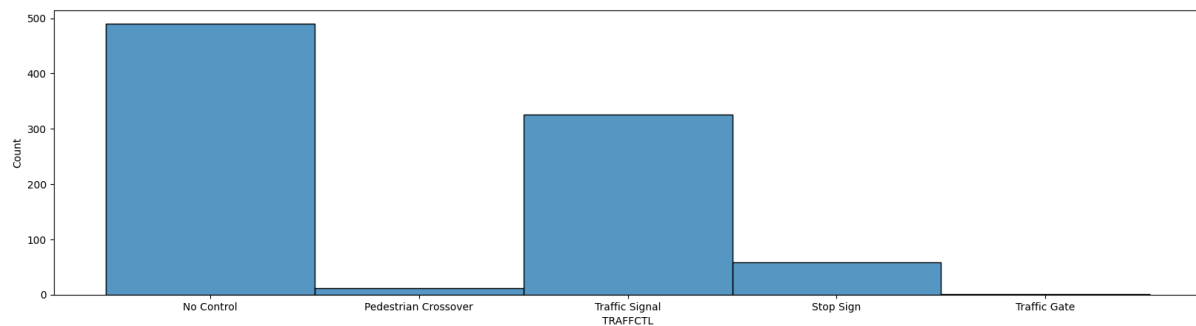
Daylight has more than 400 accidents.

We worked on binning the age groups of drivers and identified a trend in the driver which was pretty obvious.



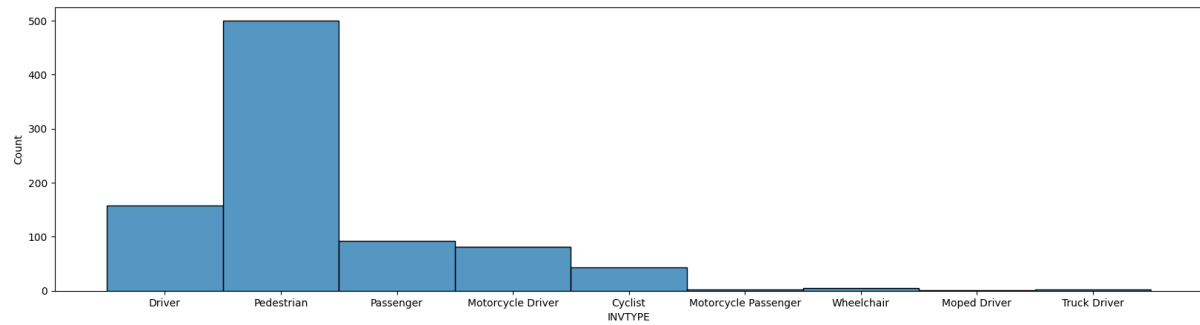
Accidents happen all Ages!

There is a category that was traffic control type or TrafficCTL



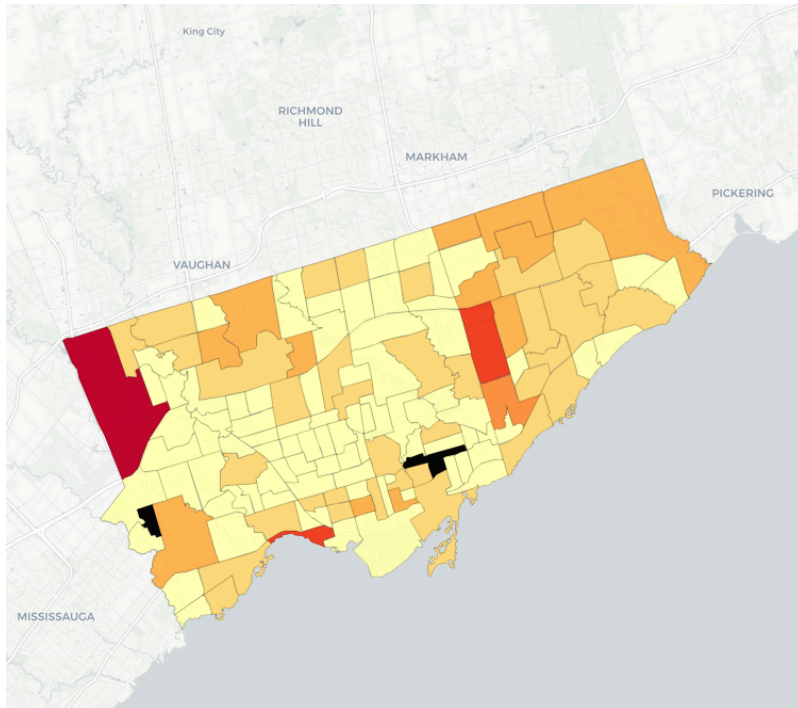
And rightfully No control was evident in most fatal collisions

People involved in the fatal collision were also identified. And most fatal collisions involved pedestrians.



Pedestrian most involved in fatal accidents

Apart from this, we also created a choropleth map to identify the most common neighborhoods for the accidents.



Highest Accidents area