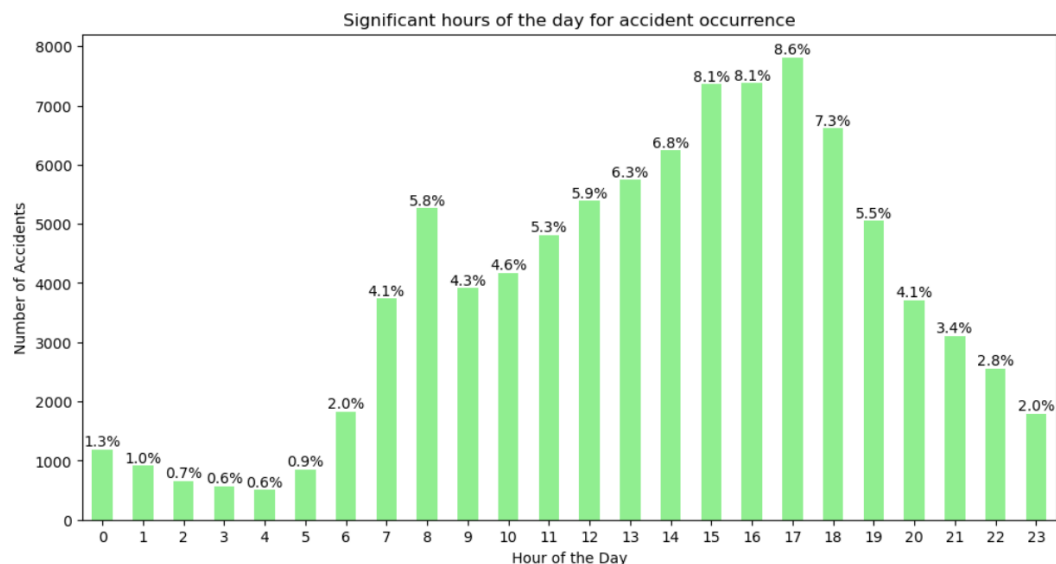


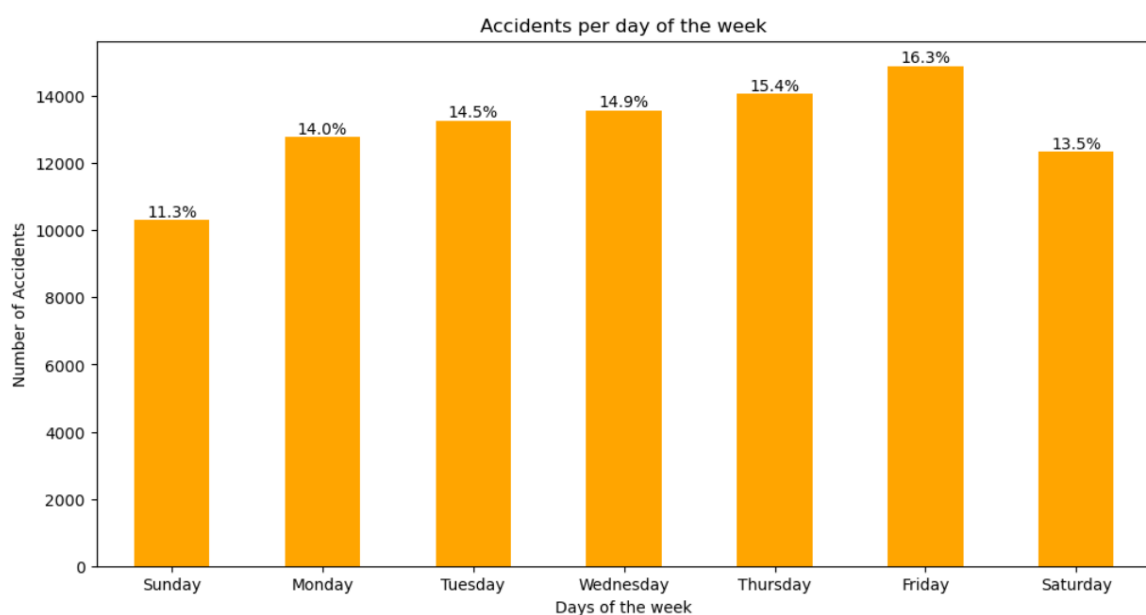
Data Insights

Significant hours of the day and days of the week for accidents occurrence.

The analysis reveals that the accidents occur throughout the day with notable trends observed during certain hours. A peak of 5.8% of accidents is observed in the morning around 8 AM. But a higher percentage of accidents have been observed during the hours between 3-5 PM reach its peak at 5 PM. The gradual rise in accidents from 10 AM to its peak at 5 PM indicates peak traffic hours as people commute to and from work



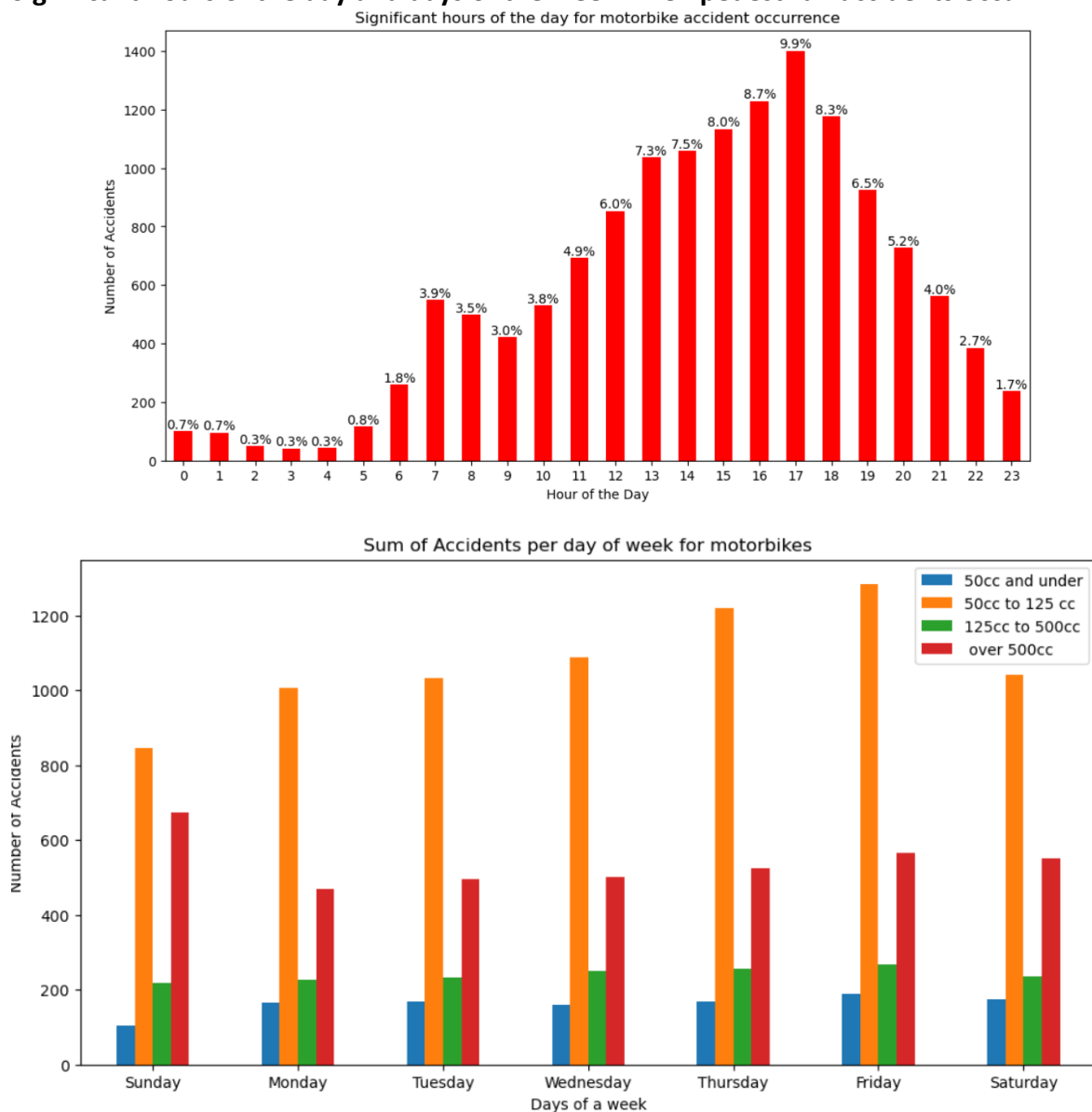
Majority of accidents occur during weekdays, with highest accident occurrence approximately at 16.3% on Fridays. On the other hand, Sundays have a lower proportion of accidents about 11.3%. The spike on Friday resulted from increase in vehicle activity due to people rushing home for the weekend.



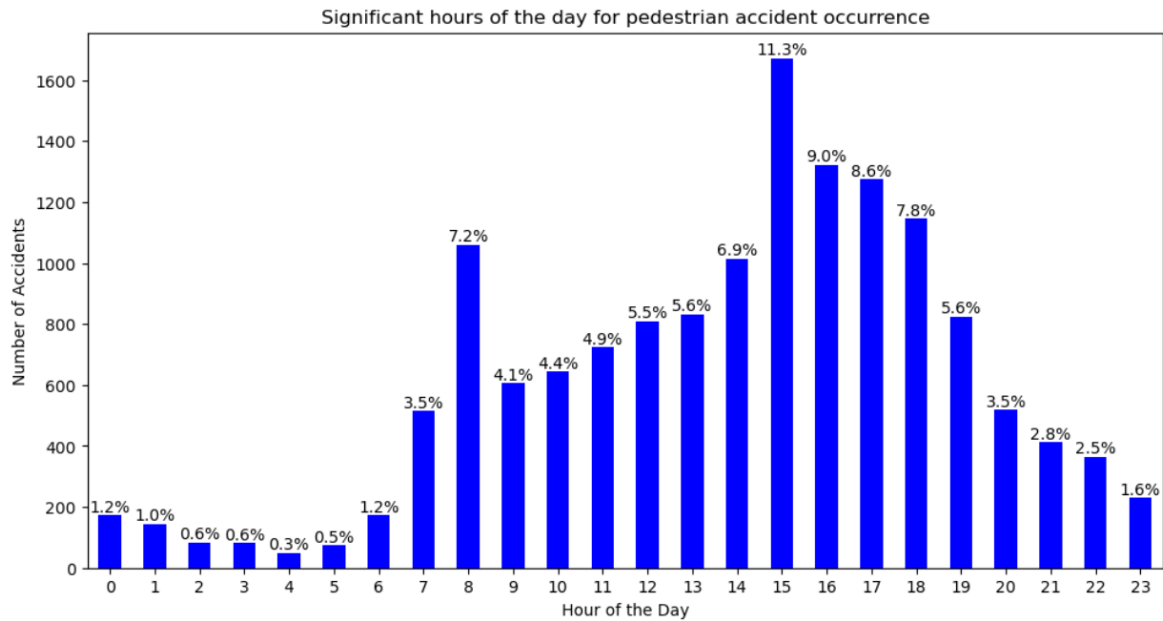
Significant hours of the day and days of the week for accidents occurrence for motorbikes.

Most of the motorbike accidents happened between 2 PM and 6 PM, with a significant peak of 9.9% total accidents occurring around 4 PM (Fig. 8). Even though, there are different categories of motorbikes, those with 50cc to 125cc engines have the highest accident rates. In terms of days of the week, Fridays are the ones that have the highest accident occurrences for 50cc and under, 50cc to 125 cc, and 125cc to 500cc engine motorbikes. However, the day with the highest accident rate that corresponds to motorbikes with 500cc engines are Sundays.

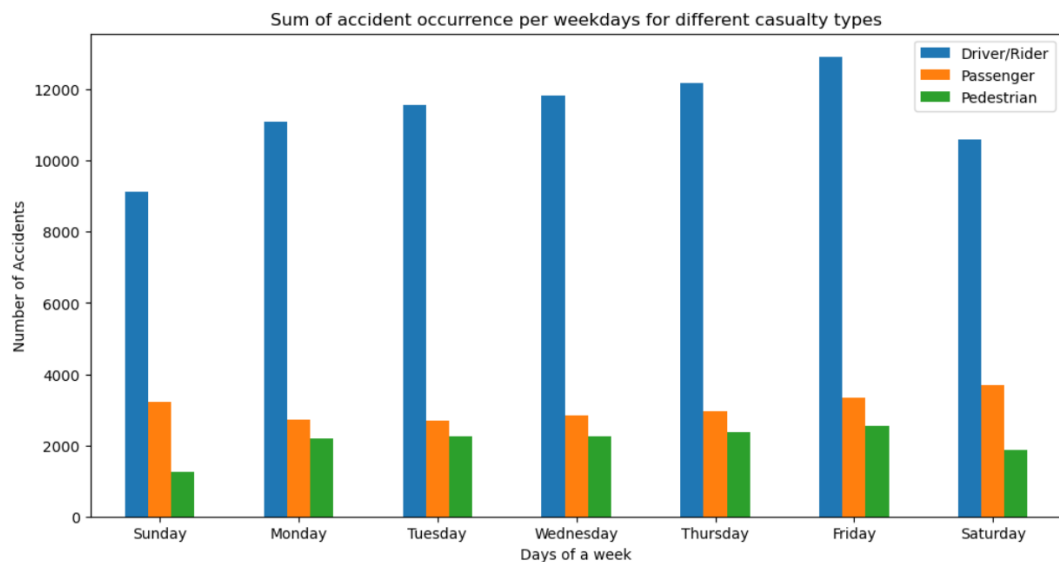
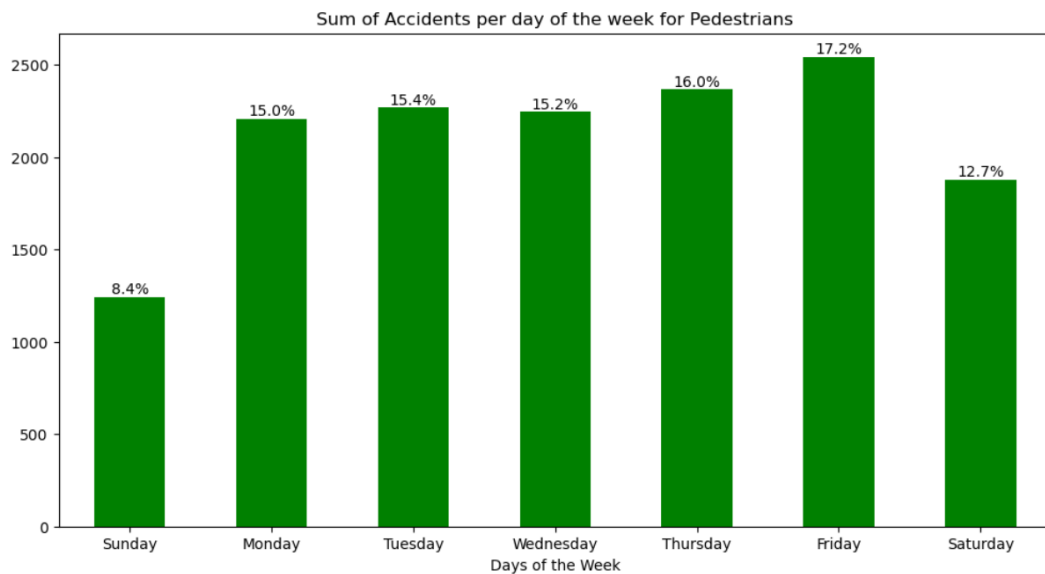
Significant hours of the day and days of the week when pedestrian accidents occur.



Higher proportion of pedestrian accidents happened between 3 PM and 6PM of a day, with a peak of 11.3% of total accidents around 3 PM. Furthermore, during the early rush hours, there were notable pedestrian accidents around 8 AM.

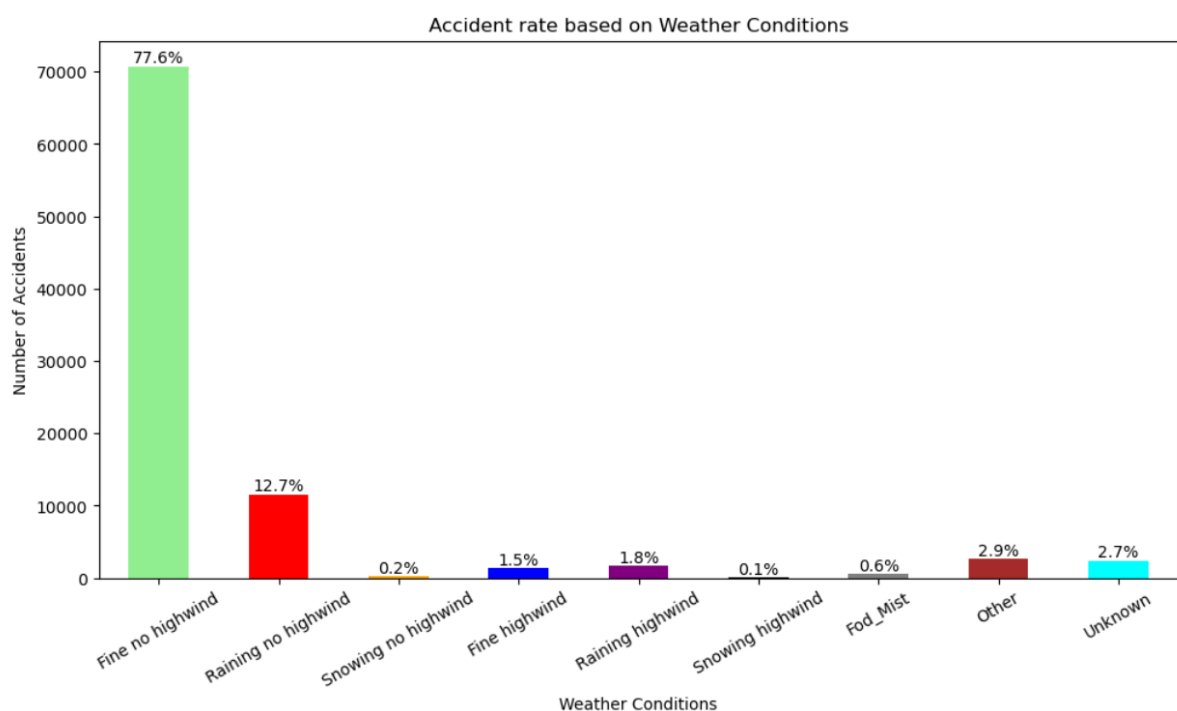
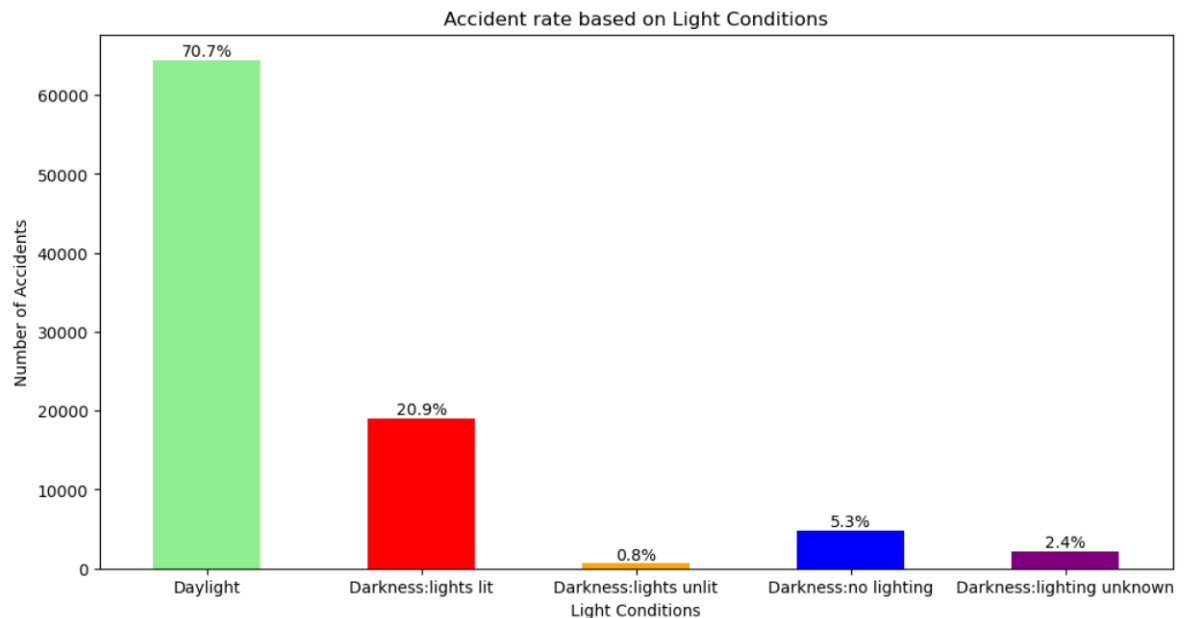


The notable decrease in accidents after 6 PM indicates that most pedestrian accidents occur during the rush hours. Additionally, with a higher peak observed on Fridays, most of the pedestrian accidents are happened throughout the weekdays (Fig. 11). Every day, a higher number of accidents involving drivers or riders occur than that of pedestrian accidents.



Effect of light and weather condition.

Interestingly, accidents aren't significantly impacted by severe weather. Over 70% of accidents happen during the day in fine weather conditions with no high winds. Darkness and poor weather often lead to a decrease the overall number of accidents as make driving undesirable, resulting in fewer vehicles on the road.



Exploring the impact of selected variables on accident severity using the apriori algorithm.

The selected features for exploring their impact on accidents using the apriori algorithm include speed limit, weather conditions, and urban or rural areas. The apriori algorithm was applied with minimum support of 0.3, minimum confidence of 0.7, and minimum lift of 1.05.

The table 1 represents the best rules that illustrate the relationship between various factors and the slight accident occurrence. For instance, with a support of 35.97%, confidence of 79.91%, and a lift of 1.02, Rule 1 suggests that a speed limit of 30 mph is associated with a higher chance of minor accidents. This indicates that minor injuries are more likely to occur in accidents that occur at slower speeds.

Rule	Rule Body	Support	Confidence	Lift
1	{Speed 30} -> {Slight}	0.359697	0.799084	1.019911
2	{Fine weather no wind, Speed 30} -> {Slight}	0.359697	0.799084	1.019911
3	{Fine weather no wind, Urban} -> {Slight}	0.359697	0.799084	1.019911
4	{Urban, Speed 30} -> {Slight}	0.459983	0.802717	1.024548
5	{Fine weather no wind, Urban, Speed 30} -> {Slight}	0.359697	0.799084	1.019911

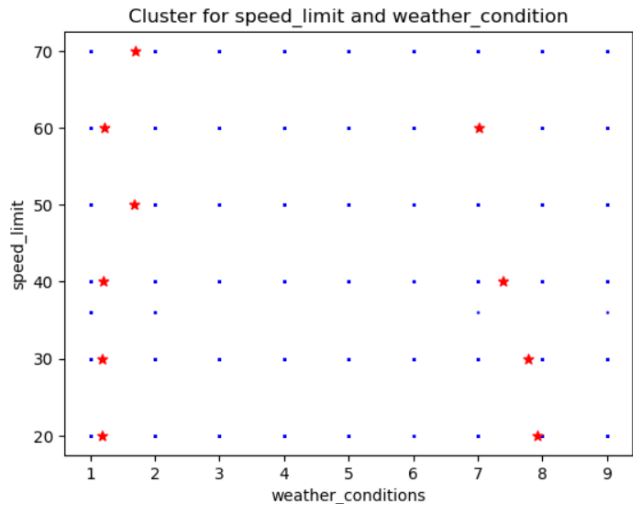
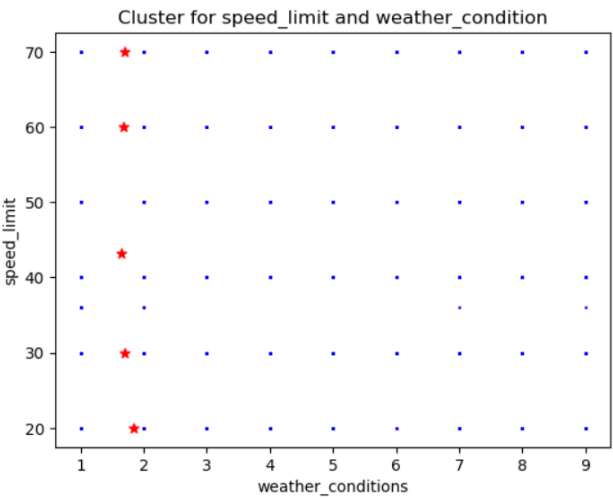
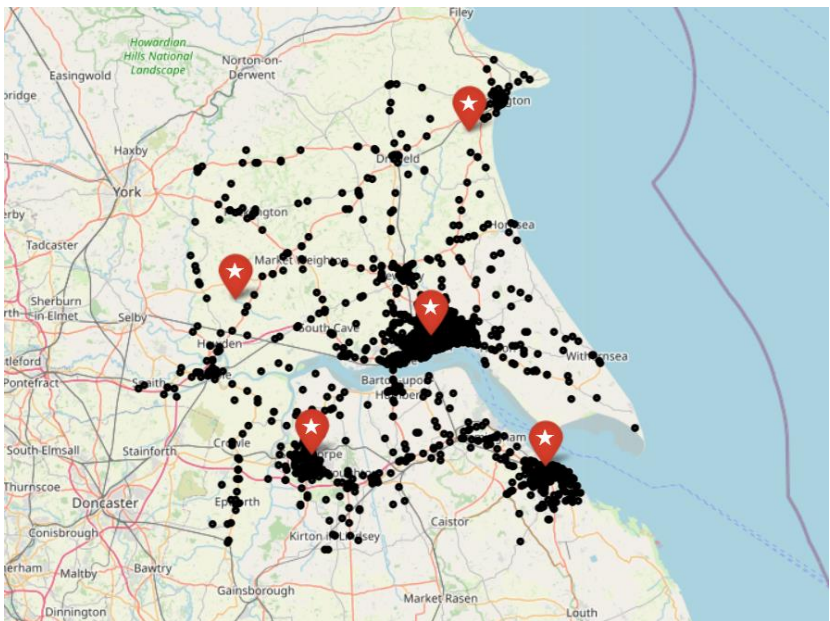
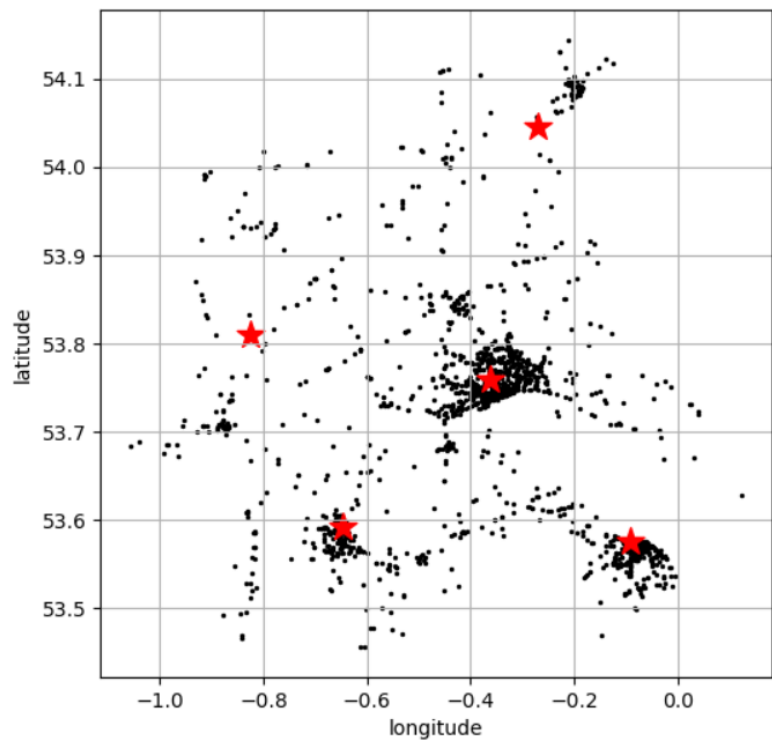
Rule 2 implies an association between fine weather conditions with no wind, a speed limit of 30 mph, and slight accidents, emphasising the effect of lower-speed environments on minor accidents. The influence of the urban environment on accident severity is further emphasised through Rule 3 that minor accidents are more frequent in urban areas, especially those with a 30-mph speed limit. These findings suggest that multiple combinations of speed limits, weather conditions, and urban environments may have an impact on the frequency of minor accidents.

Clustering Humberside region.

The distribution of accidents in the Humberside region was examined using K-Means clustering algorithm for clustering the longitude and latitude data. The optimal number of clusters was determined using the elbow curve resulting 5 clusters. From the clusters, it's evident that most of the accidents occur in nearby areas of major Humberside cities, including Grimsby, Scunthorpe, Bridlington, and Hull.

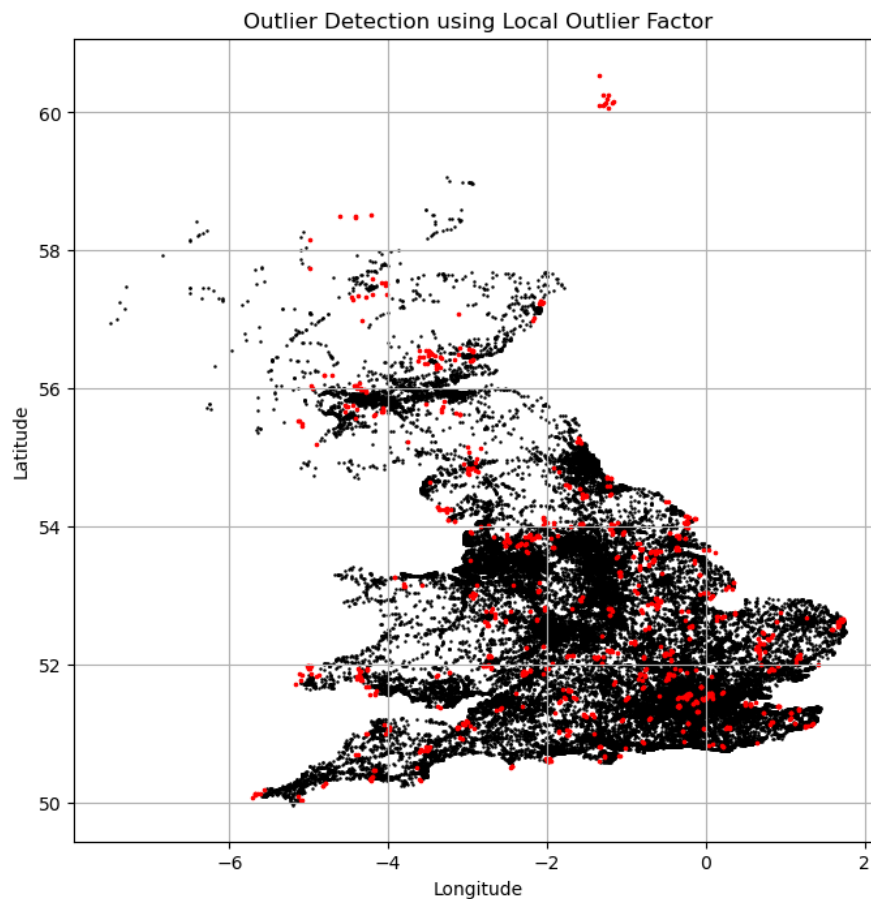
In addition to that, clustering was also carried out based on weather conditions and speed limit. The clusters were associated with weather condition 2 with number of clusters 5,

indicating raining without high winds. While increasing the number of clusters from 5 to 10 revealed insights about the impact of fog and other weather-related factors on accidents .



Outlier Detection.

The outlier detection was accomplished by employing the Local Outlier Factor (LOF) and Interquartile Range (IQR) methods. The analysis revealed that outliers have been distributed throughout the dataset, with minor concentration in and near the urban areas. These outliers may be attributed to certain circumstances.



Upon analysing outliers associated with the Humberside region using the IQR method, the results revealed that the outliers are not closely located to the major cities in the Humberside region. It was decided to retain the outliers in the dataset since these outliers might be correlated with specific circumstances or scenarios.

