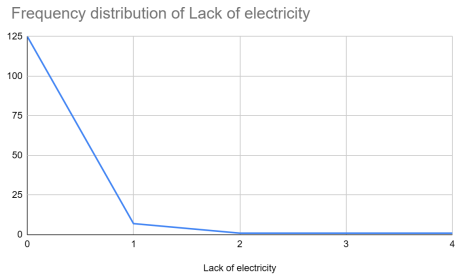
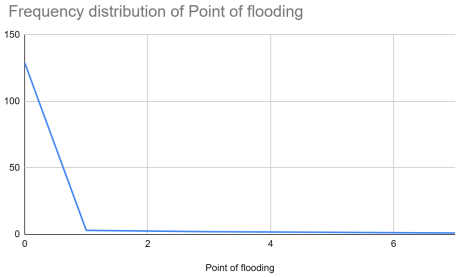
1. High Impacting Variables: From our analysis, we have seen that point of flooding and lack of electricity greatly impacts our congestion rate.

Let's closely analyze their distribution.

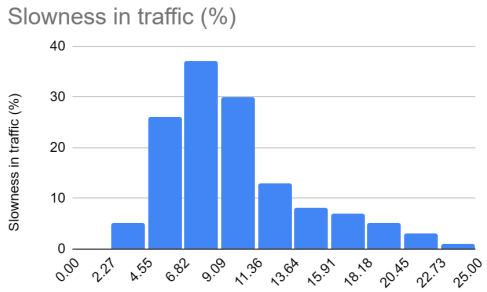


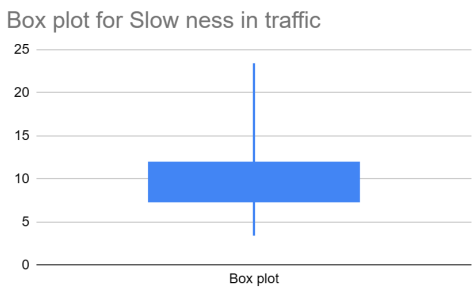


These 2 variables have high frequencies of 0 which clearly defines their absence. This means that they are able to create disturbances even with small frequencies of events. Therefore, these variables require our immediate and quick attention.

Now said that, there may not always be a direct causal relationship between these variables to congestion rates which may require more domain information.

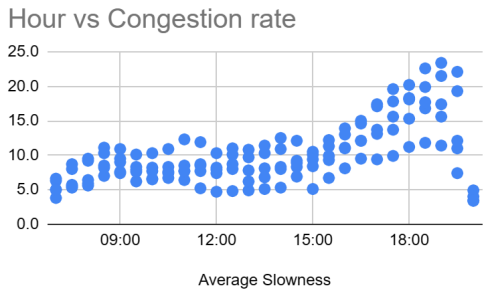
2. High Congestion Rates: Though most of the congestion rate centers around 9-10%. There are values that go as high as 25% which clearly indicates the added on problems in the form of existing variables present in the data. Rather than removing these extremes from the data, we need to act swiftly on them.

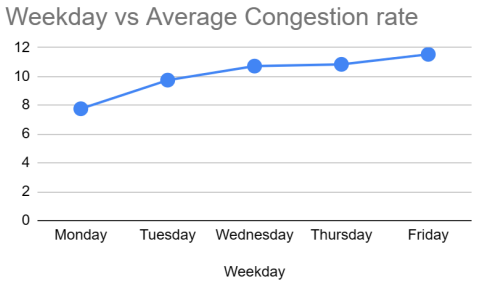




3. Peak Hours: We have witnessed in our dataset that timing has a huge impact on the slowness in traffic. This is one of the most important variables which has a positive relationship with the congestion rate.

> These peak hours are critical, and there is a risk of high congestion rate during these times, especially on weekdays. To limit this risk, we can resume our operations on the road network within non-peak hours..





By addressing these challenges and capitalizing on the insights drawn from the data, we believe the organization can significantly enhance citizen's satisfaction as well as operational efficiency.