

ABSTRACT

NY City's ECB (Environmental Control Board) NOTICE OF VIOLATIONS and PARKING VIOLATIONS datasets are used where I am going to analyze which area is more prone to both the violations at a particular interval of time (like is it during day time, evening time or night time). Also I would like to merge this analysis with weather data where I will try to find out a pattern like in what kind of weather conditions that particular area is more prone to violations. For this analysis I am going to use the clustering algorithm. **Cluster analysis** or **clustering** is the task of grouping a set of objects in such a way that objects in the same group (called a **cluster**) are more similar (in some sense or another) to each other than to those in other groups (clusters).