

# Exploratory Data Analysis

## Key Findings

During exploratory data analysis, we ask the following questions:

1. What kind of factors affect the safety level of childcare facilities?
2. Is there a relationship between the inspection report findings and the ratings of the parents?
3. Does location/neighborhood affect the safety level of childcare centers?

Initial findings are:

- As we expected, the number of schools are correlated with the number of violations. Brooklyn is the region, which has the highest number of schools in the New York City area, and Staten Island has the least number of schools. The highest number of violation has been observed in the Brooklyn region and the number of observed violations in Staten Island is the smallest.
- The number of group daycare is higher than other types of facilities in each region.
- 3698 inspection visits turn out without any violation observation.
- The number of inspections visit that violation observed is 38081.
- The maximum number of inspection visit to a facility is 74. It means approximately 24 visit in a year and 2 visit in a month.
- There isn't a correlation between the number of educated staff of a facility and the number of the violation. However, many parents trust facilities more if educated staff number is higher.
- There is a relation between the number of years that the facility operated and the number of the violation. The longer time operated facilities have higher violation numbers.
- There isn't any educated staff in camps.
- The highest number of violations observed between 2-5 year-old age ranges.

- The number of violations is least in summer camps. However, summer camps are short term facilities. They operate for three months. It could affect the amount of visit for inspection.
- There isn't a correlation between the number of violations and the ratings at Yelp. It means parents are not a good detector for violations.
- There are some neighborhoods that have facilities only with high number of violations.
- Number of general violation is 12261, number of public health hazard violation is 4319, and number of critical violations is 10318.
- Number of uncorrected violation status is 1262.
- The worst 3 child care centers are:

**THE LEARNING TREE 74 inspection visit with violations**

**TENDER TOTS/137 ST LLC 73 inspection visit with violations**

**SEABURY DAY CARE CENTER 73 inspection visit with violations**