Methodology:

The dataset covers results between 2020-2023 from Department of Health & Mental Hygiene Inspections of New York City school cafeteria and looks at a variety of health code violations.

This data was analyzed by creating a pivot table to identify the most common violation types and where they occurred. Upon analyzing the pivot table results, it was determined that the most common violations are 10F, 08A, and 04L. 08A violations refer to an establishment that "is not free of harborage or conditions conducive to rodents, insects or other pests." 04L violations refer to "evidence of mice or live mice in food or non-food areas." 10F violations refer to "Non-food contact surface or equipment made of unacceptable material, not kept clean, or not properly sealed, raised, spaced or movable to allow accessibility for cleaning on all sides, above and underneath the unit." All as defined by the data dictionary.

Initially, the goal was to conduct a historical analysis to see if rodent violations had gone down since the "war on rats" began, but then found that the dataset was not complete enough for an accurate historical analysis. Some schools are able to appeal their violation status and therefore schools may be removed from the list if a ruling is made in their favor. It also only includes data from active school cafeterias, so if a school cafeteria closes down, their data is removed from the dataset. An analysis to look at the most common violations for 2023 was then conducted. The data was then by years. This was done by splitting the rows by a slash to separate years, dates and months. All of the available data for 2023 was then moved into a new sheet for analysis with a pivot table. The spreadsheet with my cleaned data used for analysis can be found here.