

## Description of data set

The dataset contains every sustained or not yet adjudicated violation citation from every full or special program inspection conducted up to three years prior to the most recent inspection for restaurants and college cafeterias in an active status on the RECORD DATE (date of the data pull). When an inspection results in more than one violation, values for associated fields are repeated for each additional violation record. Establishments are uniquely identified by their CAMIS (record ID) number. Keep in mind that thousands of restaurants start business and go out of business every year; only restaurants in an active status are included

In the dataset.

Records are also included for each restaurant that has applied for a permit but has not yet been inspected and for inspections resulting in no violations. Establishments with inspection date of 1/1/1900 are new establishments that have not yet received an inspection. Restaurants that received no violations are represented by a single row and coded as having no violations using the ACTION field. Because this dataset is compiled from several large administrative data systems, it contains some illogical values that could be a result of data entry or transfer errors. Data may also be missing. This dataset and the information on the Health Department's Restaurant Grading website come from the same data source. The Health Department's Restaurant Grading website is here: <http://www1.nyc.gov/site/doh/services/restaurant-grades.page>

Column Name	Description	Type
CAMIS	This is an unique identifier for the entity (restaurant); 10-digit integer, static per restaurant permit	Plain Text
DBA	This field represents the name (doing business as) of the entity (restaurant); Public business name, may change at discretion of restaurant owner	Plain Text
BORO	Borough in which the entity (restaurant) is located.; • 1 = MANHATTAN • 2 = BRONX • 3 = BROOKLYN • 4 = QUEENS • 5 = STATEN ISLAND • Missing; NOTE: There may be discrepancies between zip code and listed boro due to differences in an establishment's mailing address and physical location	Plain Text
BUILDING	Building number for establishment (restaurant) location	Plain Text
STREET	Street name for establishment (restaurant) location	Plain Text
ZIPCODE	Zip code of establishment (restaurant) location	Plain Text

Column Name	Description	Type
PHONE	Phone Number; Phone number provided by restaurant owner/manager	Plain Text
CUISINE DESCRIPTION	This field describes the entity (restaurant) cuisine. ; Optional field provided by provided by restaurant owner/manager	Plain Text
INSPECTION DATE	This field represents the date of inspection; NOTE: Inspection dates of 1/1/1900 mean an establishment has not yet had an inspection	Date & Time
ACTION	This field represents the actions that is associated with each restaurant inspection. ; • Violations were cited in the following area(s). • No violations were recorded at the time of this inspection. • Establishment re-opened by DOHMH • Establishment re-closed by DOHMH • Establishment Closed by DOHMH. Violations were cited in the following area(s) and those requiring immediate action were addressed. • "Missing" = not yet inspected;	Plain Text
VIOLATION CODE	Violation code associated with an establishment (restaurant) inspection	Plain Text
VIOLATION DESCRIPTION	Violation description associated with an establishment (restaurant) inspection	Plain Text
CRITICAL FLAG	Indicator of critical violation; "• Critical • Not Critical • Not Applicable"; Critical violations are those most likely to contribute to food-borne illness	Plain Text
SCORE	Total score for a particular inspection; Scores are updated based on adjudication results	Number
GRADE	Grade associated with the inspection; • N = Not Yet Graded • A = Grade A • B = Grade B • C = Grade C • Z = Grade Pending • P= Grade Pending issued on re-opening following an initial inspection that resulted in a closure	Plain Text
GRADE DATE	The date when the current grade was issued to the entity (restaurant)	Date & Time
RECORD DATE	The date when the extract was run to produce this data set	Date & Time
INSPECTION TYPE	A combination of the inspection program and the type of inspection performed; See Data Dictionary for full list of expected values	Plain Text
Latitude		Number
Longitude		Number

Column Name	Description	Type
Community Board		Plain Text
Council District		Plain Text
Census Tract		Plain Text
BIN		Plain Text
BBL		Plain Text
NTA		Plain Text

We will need reliable location data from locations in different neighborhoods in the Brooklyn borough. In order to obtain this information, we will use the "Foursquare" location information. Foursquare is a location data provider with information on all kinds of locations and events in an area of interest. This information includes place names, locations, menus and even photos. After finding the list of neighborhoods, we then log on to the Foursquare API to collect information about the locations in each neighborhood. For each neighbourhood, we chose the 100-metre radius The data extracted from Foursquare contained information about the sites at a specified distance from the longitude and latitude of the postcodes. The information obtained by site as follows:

- Neighborhood
- Neighborhood Latitude
- Neighborhood Longitude
- Venue
- Name of the venue e.g. the name of a store or restaurant
- Venue Latitude
- Venue Longitude
- Venue Category