

# About Restaurant Inspection Data from NYC Open Data

---

## Contents

Introductory Notes.....	1
About the Records in the Dataset.....	2
Updates to Data Set.....	2
Understanding the Letter Grading Program .....	2
Identifying Gradable Inspections .....	3
Understanding Adjudication .....	4
Examples: Calculate Most Recent Grade by Restaurant.....	5
Example SQL query to identify most recent letter grade by restaurant.....	6
Example R script to identify most recent letter grade by restaurant .....	7
Additional Resources .....	9

## Introductory Notes

This dataset includes NYC restaurant and college cafeteria (hereafter, restaurants) inspection results for up to three years prior to the most recent inspection. The purpose is to provide information on recent inspection results. Restaurants that go out of business are removed. Therefore, this dataset is not appropriate for historical analyses of NYC restaurant inspections that compare previous years of data to the current data. In addition, restaurants can choose to go through the adjudication process, i.e., argue their case at an administrative hearing. Restaurants also have appeal rights, and the entire adjudication process from start to finish can take several months. Scores current as of today may be revised due to adjudication in subsequent weeks or months. The change in scores due to adjudication is another reason why it is not valid to compare current scores to scores from previous years.

**The analytic guidance provided in this document represents the very minimum of what needs to be understood to properly use and interpret the restaurant inspection data. Please read this document in its entirety before using the data.**

**Letter grading inspections were put on pause beginning March 17, 2020, until July 19, 2021, due to the COVID-19 public health emergency. Modified restaurant inspections occurred during this time.**

## About the Records in the Dataset

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>

See the data dictionary file in the Attachments section of the OpenData website for a summary of data fields and allowable values.

## Updates to Data Set

The dataset currently available on OpenData differs from previous versions in that administrative and other unscored violations are included. We have also recently added fields that describe INSPECTION TYPE and GRADE associated with inspections.

## Understanding the Letter Grading Program

The analyst must be familiar with the NYC restaurant letter grading program to be able to analyze this dataset correctly. Background on how the program works can be found on the DOHMH website:

<https://www1.nyc.gov/site/doh/business/food-operators/letter-grading-for-restaurants.page>

Additional information, including a full list of violations, can be accessed here:

<http://www1.nyc.gov/assets/doh/downloads/pdf/rri/blue-book.pdf>

In brief, restaurant letter grading began July 27, 2010. The program allows for a two-step inspection process, providing an opportunity for restaurants who do not receive an "A" on their initial inspection to be re-inspected. This re-inspection occurs no less than 7 days after the initial inspection.

- A score of less than 14 points on either initial or re-inspection results in an “A” grade
- On re-inspection, a score of 14-27 points means a restaurant receives both a “B” grade and a “Grade Pending” card
- On re-inspection, a score of 28 or more points means a restaurant receives both a “C” grade and a “Grade Pending” card
- Certain uncorrected violations may generate additional compliance inspections, which are ungraded.

The restaurant is instructed to post the grade card or the grade pending card, but some restaurants may post both. Adjudication of a notice of violation at the Office of Administrative Trials and Hearings (OATH) – Hearings Division may result in a grade change and the issuance of a new grade card. Depending on the result of adjudication, as of the decision date, a restaurant must post the grade card issued at the time of the inspection – if violations were sustained, or it must post the grade determined by OATH-Hearings Division. Restaurants currently also have the option of foregoing a hearing and accepting a settlement offer, which represents an admission of all violations. Once the restaurant has had a hearing or has settled, a grade is considered post-adjudicated. The next inspection cycle is determined by the higher of the initial and re-inspection pre-adjudicated scores. If the initial score is less than 14 points, the next initial inspection occurs approximately 12 months later; if the higher of the scores is 14-27 points, the next initial will be 5 to 7 months after re-inspection; if the higher score is 28 or more points, the next initial inspection will occur 3 to 5 months after re-inspection.

If a restaurant is closed on inspection, the process is a little different.

If a restaurant is closed on a re-inspection, it will get a reopening inspection before it is allowed to resume operations. The reopening inspection is not a gradable inspection and therefore the score will not determine the grade for the inspection cycle. If the restaurant passes the reopening inspection, the CURRENT GRADE field will be based on the re-inspection score the restaurant received before it was closed.

If a restaurant is closed on an initial inspection, it will get a reopening inspection before it is allowed to resume operations. The reopening inspection is not a gradable inspection and therefore the score will not determine the grade for the inspection cycle. If the restaurant passes the reopening inspection, the previously posted grade card is taken down and it is given a “Grade Pending” card until a re-inspection determines the next grade.

## Identifying Gradable Inspections

Not every inspection is a “gradable” inspection. Gradable inspections can be identified when the following statements are true:

INSPECTION TYPE in (Cycle Inspection/Initial Inspection, Cycle Inspection/Re-Inspection, Pre-Permit (Operational)/Initial Inspection, Pre-Permit (Operational)/Re-Inspection)

ACTION in (Violations were cited in the following area(s), No violations were recorded at the time of this inspection, Establishment Closed by DOHMH)

INSPECTION DATE >= July 27, 2010

Although this logic will identify gradable inspections, grades associated with a closing inspection are not assigned until the reopening inspection. For this reason, the score and assigned grade on a reopening inspection are not consistent.

EXAMPLE 1:

CAMIS	INSPECTION DATE	ACTION	SCORE	GRADE	INSPECTION TYPE
40377630	1/17/2014	Establishment Closed by DOHMH. Violations were cited in the following area(s) and those requiring immediate action were addressed.	43		Cycle Inspection / Re-inspection
40377630	1/21/2014	Establishment re-opened by DOHMH	2	C	Cycle Inspection / Reopening Inspection

The Health Department conducts non-gradable inspections to enforce Special Programs, such as calorie labeling or the Smoke Free Air Act, or they may be performed for other reasons, such as when a restaurant is ready to reopen after a closing. Some of these inspections are conducted when a restaurant has not yet started or resumed operations, and these non-operational inspections tend to receive lower scores since the restaurant is not in operation during the inspection.

Special Program and Administrative violations (including violation codes 15 - 22) do not contribute towards SCORE and do not contribute towards determining GRADE. Compliance inspections are scored but do not contribute towards determining GRADE.

## Understanding Adjudication

The SCORE field contains post-adjudicated data when adjudication has occurred as of the RECORD DATE; data are pre-adjudicated when a restaurant is still waiting for adjudication as of this date. Note that some restaurants do not seek adjudication; they post the grade they are given at the time of inspection. If a restaurant receives a pre-adjudicated B or C grade on re-inspection, the Open Data dataset cannot distinguish whether that restaurant is posting the B or C grade, or – alternatively – the Grade Pending card that is simultaneously given.

Note that if a violation is dismissed on adjudication, it is removed from the dataset. Thus, over time, some violations will be removed in subsequent iterations of the dataset.

## DISCREPANCIES BETWEEN SCORES AND GRADES

The SCORE and GRADE fields may be inconsistent with each other because of limitations or errors in the data systems. That is to say, scores of 0-13, 14-27 and 28+ are not always accompanied by A, B and C grades, respectively, when they should be. There may also be cases where a grade card was given out but a record of that grade issuance is missing from the data system, and therefore missing from this dataset, even though the SCORE field is populated. Note that when initial inspections are adjudicated down to the A range, the absence of an accompanying grade associated with that inspection is correct, because the grade would not be assigned until the re-inspection is performed.

### **Examples: Calculate Most Recent Grade by Restaurant**

NYC Open Data allows access to this data via API or by downloading a copy of the current data set in a static format. Included below are SQL and R examples for calculating the most recent grade by restaurant.

## Example SQL query to identify most recent letter grade by restaurant

```
--Notes:
--Only some inspection types result in gradable inspections
--While scores are provided on initial inspection, grades are not unless the score
is less than or equal to 13. If a restaurant receives a score of > 13 on initial
inspection, the current grade will still reflect the grade received during the
last inspection cycle until the next re-inspection is complete.

-- Download DOHMH NYC Restaurant Inspection Results data set and save as CSV file:
NYC_Insp_Results.csv
```

```
with RecentInspDate as (
  select CAMIS
    , max([INSPECTION DATE]) as MostRecentInspDate
  from NYC_Insp_Results
 where ([INSPECTION TYPE] in (
    'Cycle Inspection / Re-inspection'
    , 'Pre-permit (Operational) / Re-inspection')
  OR ([INSPECTION TYPE] in (
    'Cycle Inspection / Initial Inspection'
    , 'Pre-permit (Operational) / Initial Inspection')
  AND SCORE <= 13) or ([INSPECTION TYPE] in (
    'Pre-permit (Operational) / Reopening Inspection'
    , 'Cycle Inspection / Reopening Inspection')
  )) and GRADE in ('A', 'B', 'C', 'P', 'Z') --values where a grade card or
  grade pending card is issued
 group by CAMIS)
```

```
--Select restaurant inspection data based on the most recent inspection date
select distinct r.CAMIS
  , DBA as Name
  , r.MostRecentInspDate
  , GRADE
  , [INSPECTION TYPE]
  , SCORE
from NYC_Insp_Results i
join RecentInspDate r
  on r.CAMIS = i.CAMIS
  and r.MostRecentInspDate = i.[INSPECTION DATE]
where [INSPECTION TYPE] in (
  'Cycle Inspection / Re-inspection'
  , 'Pre-permit (Operational) / Re-inspection' -- re-inspections
  , 'Pre-permit (Operational) / Reopening Inspection'
  , 'Cycle Inspection / Reopening Inspection') --re-opening inspections where
  grade pending is issued
OR ([INSPECTION TYPE] in (
  'Cycle Inspection / Initial Inspection'
  , 'Pre-permit (Operational) / Initial Inspection')
  AND SCORE <= 13) --initial inspections where A grade is issued
```

## Example R script to identify most recent letter grade by restaurant

```
#Install packages: reader, tidyverse
```

```
library(readr)
```

```
library(tidyverse)
```

```
# Download DOHMH NYC Restaurant Inspection Results data set and save as CSV file:  
NYC_Insp_Results.csv
```

```
Open_Data_Sample <- read_csv("DOHMH_New_York_City_Restaurant_Inspection_Results.csv",  
                             col_types = cols(ZIPCODE = col_character())  
                             )
```

```
View(Open_Data_Sample)
```

```
#Filter on inspection type, score, grade
```

```
Inspections <- Open_Data_Sample %>%  
  filter(('INSPECTION TYPE' %in%  
          c('Cycle Inspection / Re-inspection'  
            , 'Pre-permit (Operational) / Re-inspection')  
          | ('INSPECTION TYPE' %in%  
            c('Cycle Inspection / Initial Inspection'  
              , 'Pre-permit (Operational) / Initial Inspection'))  
          & SCORE <= 13)  
          | ('INSPECTION TYPE' %in%  
            c('Pre-permit (Operational) / Reopening Inspection'  
              , 'Cycle Inspection / Reopening Inspection'))  
          & GRADE %in% c('A', 'B', 'C', 'P', 'Z')) %>%
```

```
select(CAMIS,`INSPECTION DATE`)
```

```
#Select distinct inspections
```

```
Inspections_Distinct <- distinct(Inspections)
```

```
#Select most recent inspection date
```

```
MostRecentInsp <- Inspections_Distinct %>%
```

```
group_by(CAMIS) %>%
```

```
slice(which.max(as.Date(`INSPECTION DATE`, '%m/%d/%Y')))
```

```
#Join most recent inspection with original dataset
```

```
inner_join(Open_Data_Sample,MostRecentInsp, by = "CAMIS","INSPECTION DATE")
```

```
#Select restaurant inspection data based on most recent inspection date
```

```
Final <- Open_Data_Sample %>% inner_join(MostRecentInsp) %>%
```

```
filter((`INSPECTION TYPE` %in%
```

```
  c('Cycle Inspection / Re-inspection'
```

```
    , 'Pre-permit (Operational) / Re-inspection'
```

```
    , 'Pre-permit (Operational) / Reopening Inspection'
```

```
    , 'Cycle Inspection / Reopening Inspection')
```

```
| (`INSPECTION TYPE` %in%
```

```
  c('Cycle Inspection / Initial Inspection'
```

```
    , 'Pre-permit (Operational) / Initial Inspection'))
```

```
& SCORE <= 13)) %>%
```



```
select(CAMIS,DBA,`INSPECTION DATE`,GRADE,`INSPECTION TYPE`,SCORE)
```

```
#Select distinct restaurant inspection data
```

```
Final <- distinct(Final)
```

```
View(Final)
```

## **Additional Resources**

[www.nyc.gov/health/foodservice](http://www.nyc.gov/health/foodservice)

<http://www1.nyc.gov/assets/doh/downloads/pdf/rii/blue-book.pdf>