

DOHMH New York City Restaurant Inspection Results

URL: <https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/43nn-pn8j>

New York health departments regularly inspect businesses serving food to ensure restaurant and retail outlets are following safe food handling procedures. I found this dataset from the New York City's open data website.

According to the Center for Disease Control and Prevention, about one in six Americans get sick, 128,00 are hospitalized, and 3,000 die of food-borne diseases. Those illnesses are caused by consuming bacteria, viruses or toxins in food. It is spread when people eat contaminated food. Public health inspection records aid customers to stay away from restaurants which have poor inspection scores.

This dataset provides restaurant inspections, violations, grades and adjudication information. This dataset has 446964 entries with 18 attributes including restaurant ID, restaurant name, borough of which the restaurant is located, building number of the restaurant, street name of the restaurant, zip code, phone number, cuisine description, inspection data, action associated to the inspection, violation code, violation description, critical flag, score, grade, record date and inspection type.

Attribute Name	Attribute Explanation	Attribute Type
<i>CAMIS</i>	Restaurant ID	Int
<i>DBA</i>	Restaurant Name	String
<i>BORO</i>	Borough in which restaurant is located	String
<i>BUILDING</i>	Building code	Int
<i>STREET</i>	Street name where restaurant is located	String
<i>ZIPCODE</i>	Zip code	Int
<i>PHONE</i>	Restaurant phone number	Int
<i>CUISINE DESCRIPTION</i>	Cuisine type of the restaurant	String
<i>INSPECTION DATE</i>	Date of the inspection	Date
<i>VIOLATION CODE</i>	Violation code	String
<i>VIOLATION DESCRIPTION</i>	Description of the violation	String
<i>CRITICAL FLAG</i>	If the inspection is critical	Boolean
<i>SCORE</i>	Score of the inspection	Int
<i>GRADE</i>	Grade of inspection	Char (A,B,...)
<i>INSPECTION TYPE</i>	Type of inspection conducted	

- Some examples of violations include sanitized equipment or utensils improperly stored, surface used to prepare raw meat is also used for ready-to-eat food, evidence of mice, etc.

The main reason I chose the dataset is that it has a large amount of data entries for extracting information. Some of the attributes may be redundant like restaurant ID, street name,

zip code, phone number. However other attributes may hide away some interesting information. Since all the restaurant is given a certain type of cuisine, we can easily tell what is the most popular cuisine type among all 5 New York City boroughs. Another example would be using Boro and score attribute, one can easily find which borough has the highest average score.

The information mention above is can be find easily. However, we may find some very interesting facts if we combine one or more attribute together. With borough attribute and cuisine type, we can easily see what restaurant is more popular in certain borough. If Queens has the most Chinese restaurant, it may imply that Queens has a higher Chinese population than other boroughs. Another fun fact would be using cuisine type and restaurant name attributes we could generalize what is the most words used to name the restaurant of a certain type of cuisine. Also by state law, the public health department inspects all restaurant once per year, inspect may occurred again if complaints are received, based on the law, we can use this dataset to gain information or predict the score and grade of a restaurant in a future inspection.

I think this datasets would now yield fun facts about restaurant, it would also produce some serious information that could make a chance in people's choice.