

Serving Up Safety: Analyzing NYC Restaurant Health Grades

Project phase: Data & EDA

03/25/2025

Ngawang Choega

nc87212n@pace.edu

Class Name: Practical Data Science

Program Name: MS in Data Science

Seidenberg School of Computer Science and Information Systems Pace
university

Agenda

- Executive summary
- Project plan recap
- Data
- Exploratory data analysis

Executive summary

Problem Statement: NYC restaurants are graded based on health inspection results, but it is difficult to predict what grade a restaurant will receive ahead of time. This makes it challenging for restaurant owners and inspectors to take proactive steps to reduce violations and improve food safety.

Solution:

- The project analyzes past health inspection records to understand what factors influence restaurant grades.
- A simple predictive system is built to estimate whether a restaurant is likely to receive an A, B, or C before an inspection.
- This helps restaurants fix issues early and allows inspectors to focus on places that need the most attention.

Project plan recap

Deliverable	Due Date	Status
Data & EDA	03/25/25	Complete
Methods, Findings, and Recommendations	04/01/25	In Progress
Final Presentation	04/29/25	Not started

Data

Data

- Key details
 - Data source: NYC Open Data – Department of Health and Mental Hygiene (DOHMH) [[NYC Restaurant Inspection Results](#)]
 - Sample size: 274,740 inspection records.([more details](#))
 - Each row represents a single violation from a restaurant inspection.
 - Time period: September 24,2015 to March 17, 2025
- What's included
 - Restaurants that are active at the time of data collection.
 - Inspections that resulted in one or more violations, as well as inspections with no violations.
 - Restaurants with grades A, B, C or pending(P).([more details](#))
- What's excluded
 - New establishments that have applied for a permit but have not yet been inspected.([more](#))
- Assumption
 - The data provides a reliable sample of active NYC restaurants and reflects typical inspection and grading outcomes.

Exploratory Data Analysis

How are NYC restaurants graded?



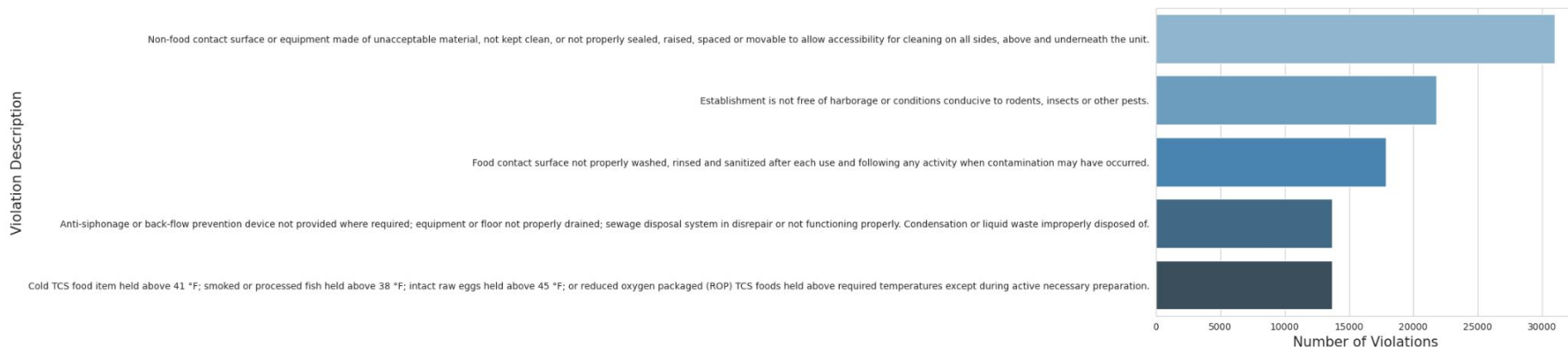
Key Takeaways

- Over 85,000 restaurants in NYC received an **A grade**, showing overall good compliance.
- A smaller portion (less than 20,000) received **B or C grades**, indicating room for improvement.
- A large number of inspections are labeled **P (Pending)** — these have been conducted, but the official grade is not yet posted.

Data notes:

- Excluded Grades: N (not yet graded) and Z (not subject to grading) were removed from the analysis.

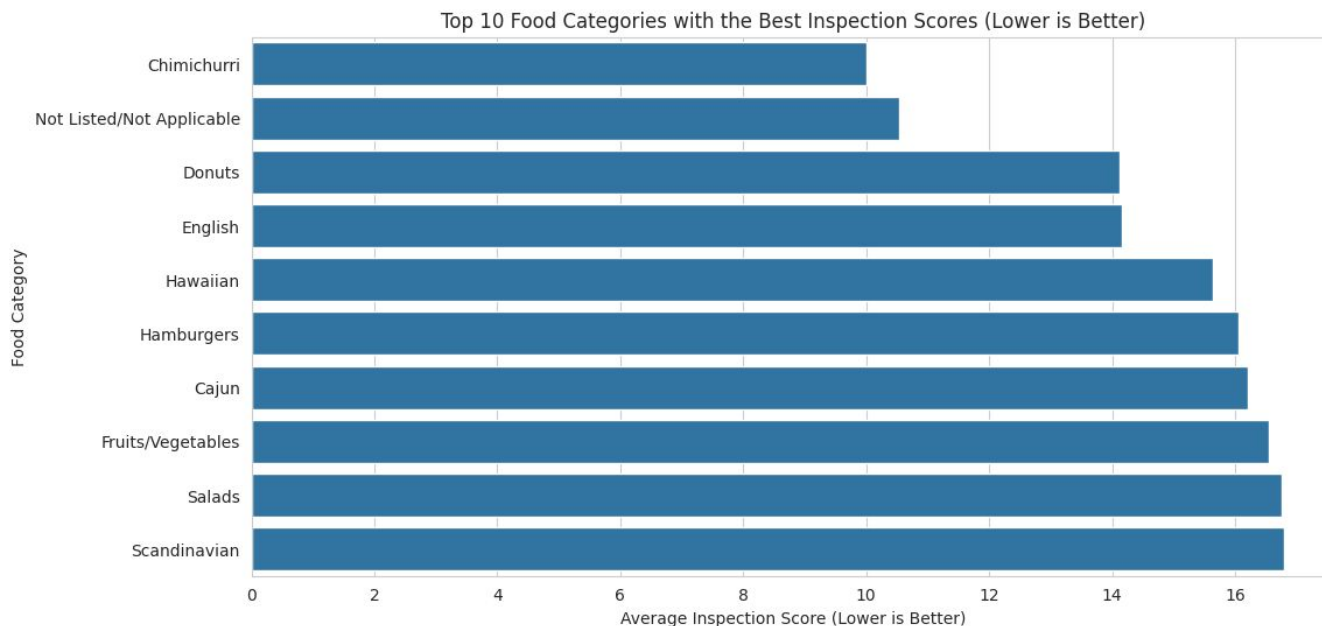
Top 5 most common violations in NYC restaurants



Key Takeaways

- The most common violation — **unclean or poorly maintained equipment** — appears in **over 30,000 inspections**.
- **Pest-related conditions** were cited in **more than 22,000 cases**, while **improper food surface sanitation** occurred in **around 18,000 inspections**.
- These high-frequency violations contribute directly to **point deductions** and are strong indicators of restaurants at risk of receiving lower grades.

Which Food Categories Receive the Best Inspection Scores?



Note: Grades are based on inspection scores. See [appendix](#) for scoring breakdown.

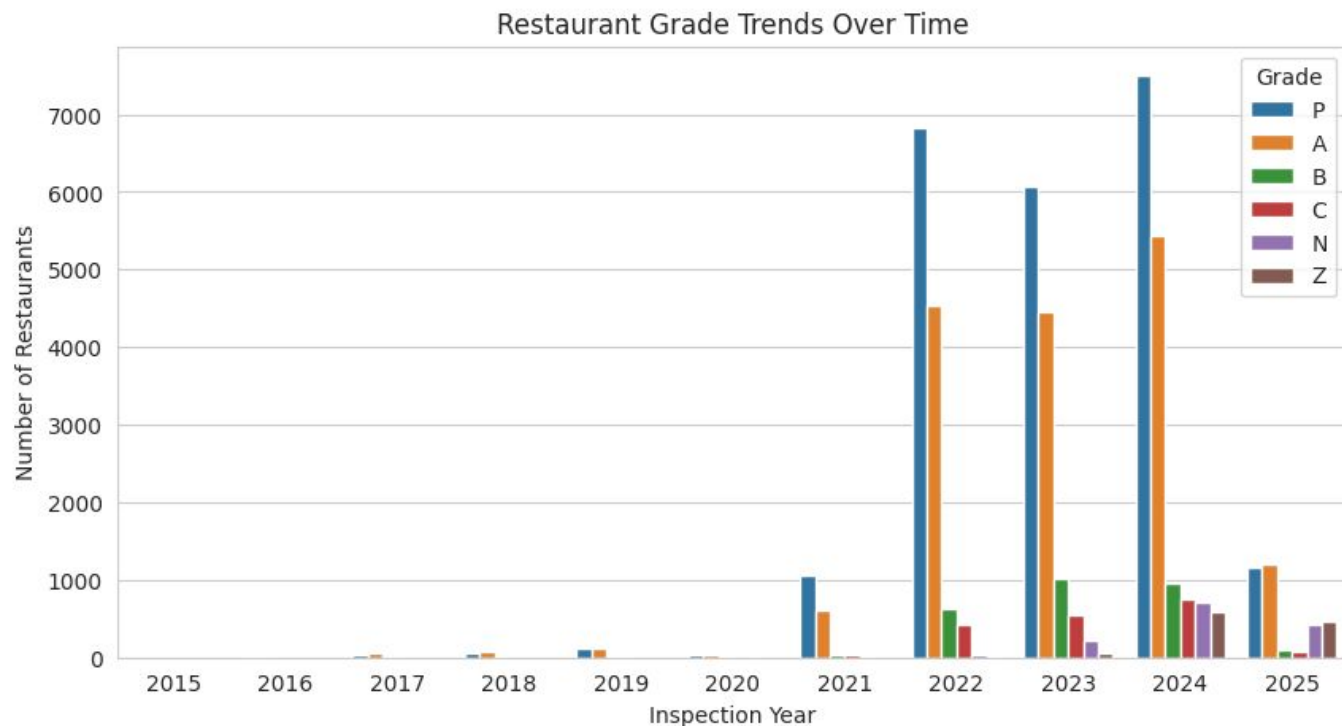
Key Takeaways

- Food categories like Chimichurri, Donuts, and English had the lowest average violation scores, suggesting better inspection outcomes.
- Scandinavian, Salads, and Fruits/Vegetables showed higher average scores, indicating more frequent issues.

Data notes:

- “Not Listed/Not Applicable” includes restaurants that did not specify a cuisine or selected a general option on their permit — often new or small businesses.

How have restaurants grade changed over time?



Note: Limited data appears from 2015–2020. See [appendix](#) for explanation.

Key Takeaways

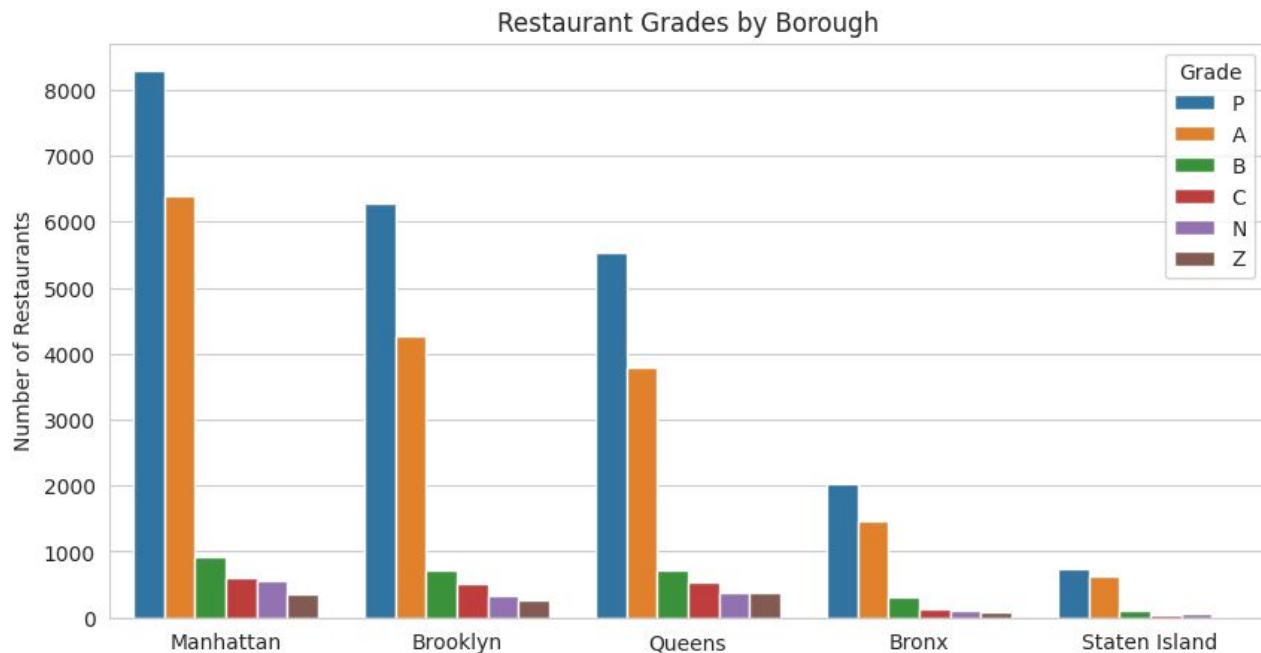
- Inspections increased sharply after 2021, following a slowdown likely caused by the pandemic.
- A grades are the most common each year, showing consistent performance.
- Many inspections from 2022 to 2024 are still marked as pending (P).

Data notes:

- N = Not yet graded (under review)
- Z = Not subject to grading (e.g., food vendors)

How do restaurant grades vary across NYC boroughs?

Key Takeaways



- Pending grades (P) are common across boroughs, especially in Manhattan (~8000+) and Brooklyn (~6000+).
- Manhattan has the highest number of restaurants, with over 6,000 A grades.
- Brooklyn and Queens follow, showing strong A grade counts with some lower grades present.
- Staten Island has the fewest restaurants (less than 1000) but the highest proportion of A grades.

Appendix

Project Materials

- Git Repo: <[link](#)>

Sample size details

- Dataset name: DOHMH_New_York_City_Restaurant_Inspection_Results.csv
- Original dataset: 274,740 rows
- Cleaned dataset: 193,892 rows
- Each row represents a single violation record from a restaurant inspection
 - If a restaurant had multiple violations during one inspection, it appears in multiple rows – one for each violation
 - If a restaurant had no violation, it is represented by a single row with a violation field marked accordingly .

Additional notes on the data

- The dataset includes records for:
 - Restaurants, college cafeterias, mobile vendors, special programs,etc.
 - New establishments that applied for a permit but have not yet been inspected (marked with the date **01/01/1900**)
 - Inspections with and without violations
- Grades used in the dataset:
 - **A, B, C** — Based on inspection outcomes
 - **P** — Grade is pending after inspection
 - **N** — Not yet graded (inspection under review or appeal)
 - **Z** — Not subject to grading (e.g., mobile vendors, special programs)



Why is there limited data from 2015-2020?

- The dataset only includes inspection records from up to three years prior to the most recent inspection for active restaurants.
- As a result, older data (pre-2021) is mostly excluded unless the restaurant had recent inspections.
- This design ensures the dataset stays focused on currently operating restaurants with recent inspection activity.

 source: [DOHMH New York City Restaurant Inspection Results](#)

NYC restaurant grading system

- A Grade: 0–13 points → Minimal violations
- B Grade: 14–27 points → Moderate violations
- C Grade: 28 or more points → Significant violations
- P Grade: Inspection done, grade pending
- N Grade: Not yet graded (under review or appeal)
- Z Grade: Not subject to grading (e.g., food vendors)

 source: [NYC restaurant grading system](#)