**Intro:**

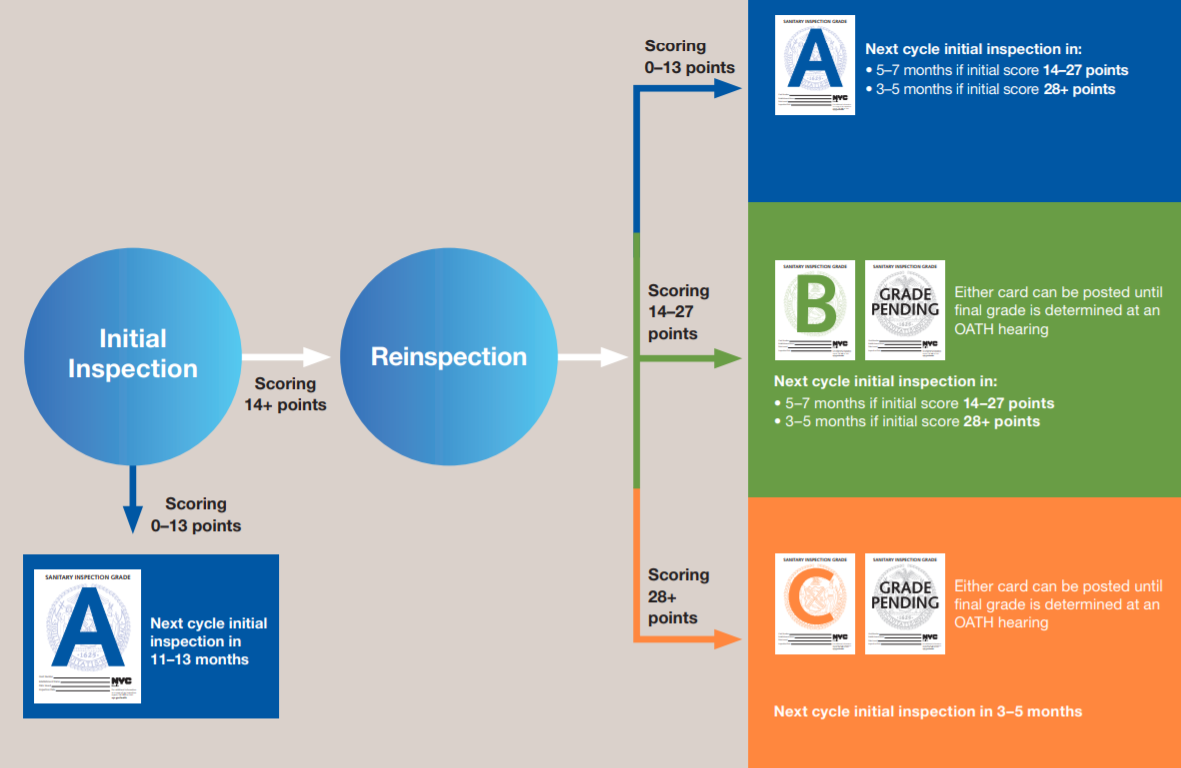
A couple years ago I ran a busy café in lower manhattan. In addition to the thin margins, constant staffing issues and downright extortionary NYC rents one is universal in the industry: that of the health inspector. Don’t get me wrong – cleanliness is essential to any restaurant owner, but when that inspector walks in it feels like judgement day. In the time since leaving that position ive had a lingerin itch to explore the NYCs open data in an effort to minimize one’s financial exposure to the DOH, while at the same time maintaining the standard of sanition that one should observe when serving (ideally) repeat guests.

The knowledge gained from this project is the advice I would observe were I to start my own food service establishment in NYC. My conclusions come solely from publicly available data as well as my own personal experience and are likely contains errors.

The inspection level data came from NYC open data (<https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/43nn-pn8j>) and since March I’ve been taking almost daily snapshots of the dataset as restaurants are constantly being deleted as the go out of business and new restaurants are opened every day. Here’s a sample of the data set:

**Timing is predictably random:**

This project started as an attempt to reverse-engineer the way the DOH selects who will be inspected on any given day. My approach was to code the logic of the inspection cycle provided in the following chart. Provided that I searched for patterns or other variables which may provide a guide.



To date I have not cracked that code. All the statistical testing I ran gave me no better than a coin flip chance. So I have no evidence to suggest that this queue selection isn’t a random generator. However, there is still value in what the results suggest. In the vast majority of cases the DOH does schedule inspections according to its stated policy. Weekend visits are extraordinarily uncommon – Sunday in particular accounts for far less than 1% percent of all visits. If, for example, you run into a case where during a shift no one has the requisite Food Handler’s Certification Card you’ll likely get away with in if it happens on the weekend. If you know which inspection bin your FSE is in you don’t have to stress out your staff unnecessarily early as this very rarely happens.

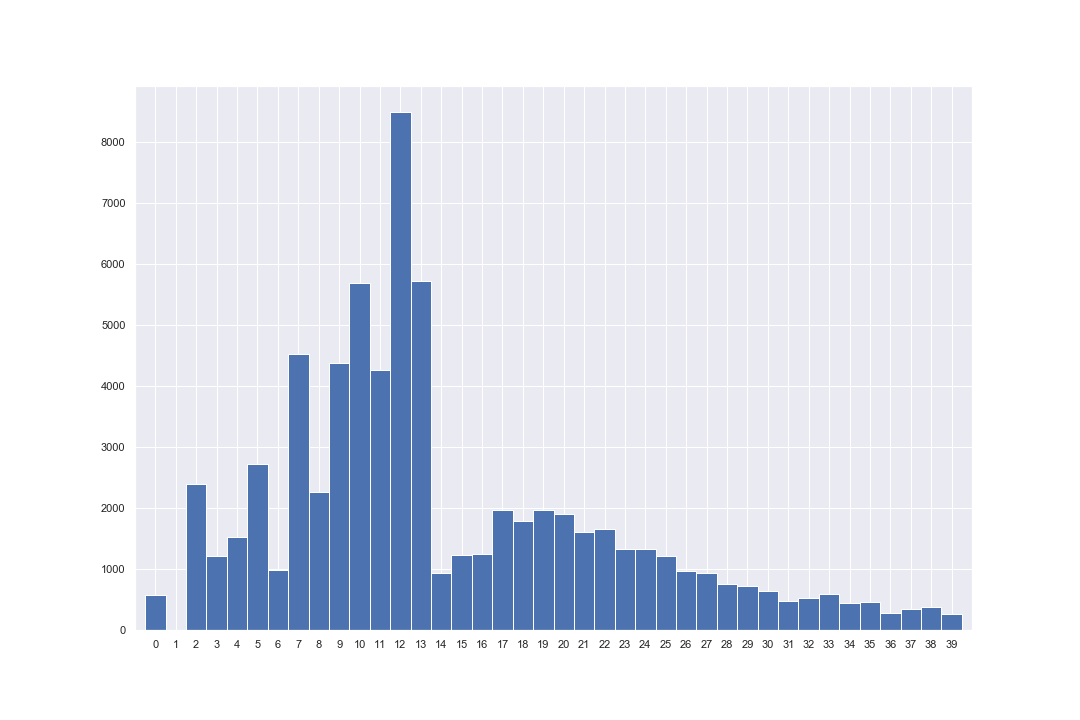
**Initial Success:**

According to health code, if a restaurant receives fewer than 14 points in violations on an initial, yearly inspection they pay no fees and the DOH leaves them alone for a whole year. At the risk of sounding obvious, all efforts should be made to achieve this. If this goal is not met it sets forth a sequence of future contact with the DOH. The following happens:

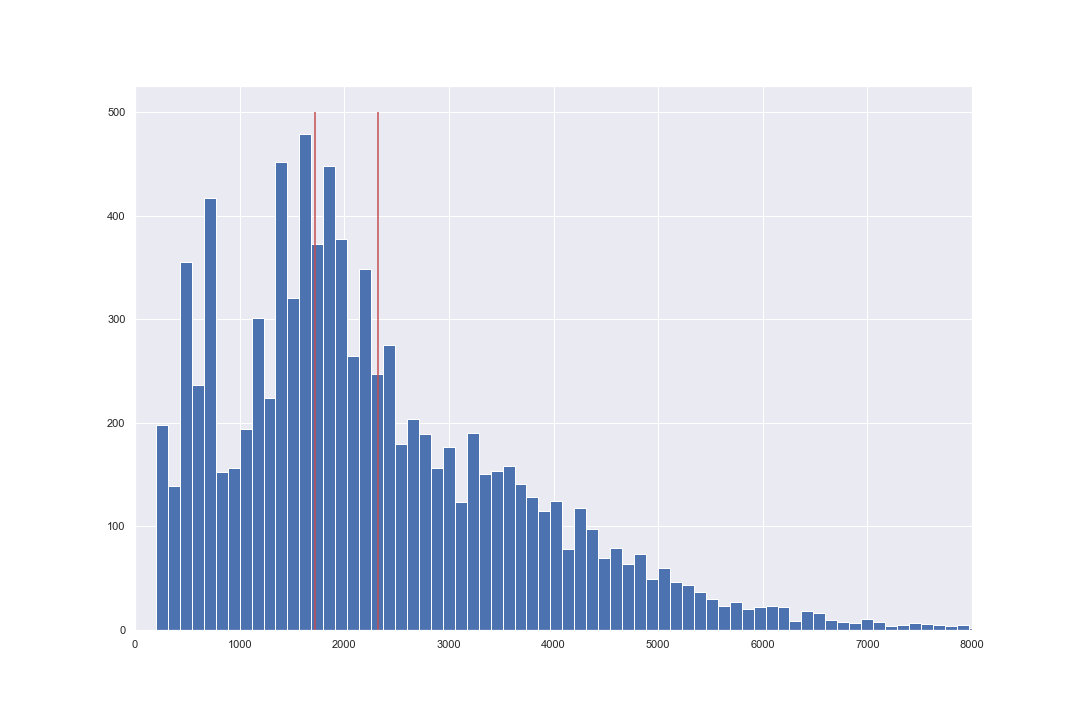
* FSE will be fined for the violations cited on initial inspection
* A reinspection will occur usually within a month and the FSE will be responsible to pay fines for violations found on this visit.
* A grade pending sign must be posted in the front window, which at best doesn’t hurt sales
* The restaurant will face another initial cycle inspection within 3 to 7 months, depending on the first score.

In this plot we see the distribution of scores on initial inspection up to 39 (the highest in the dataset is 164, yikes!). It turns out that “passing” an initial inspection isn’t so uncommon. 39% of inspections received 13 or fewer points.

Discuss how this is a strange distribution. Perhaps the result of an internal policy?



The source of my fine estimates is the NYC fine schedule. Each violation has up to five levels of severity and accompanying fines. I mapped the minimum and maximum possible fine for each inspection to arrive at my average estimates. There can be a pretty wide range between the 2 estimates but it’s something. If we look at the 2018 fine estimates conditional on “failing” the initial we see this distribution. The two red lines correspond to the middle quintile.



This information could be useful to a restaurant manager or owner when budgeting for future expenses. One can strive for perfection but can still make an educated guess that if you do have a hiccup on your inspection you can reasonably expect to cut checks to the city totaling around $2000.

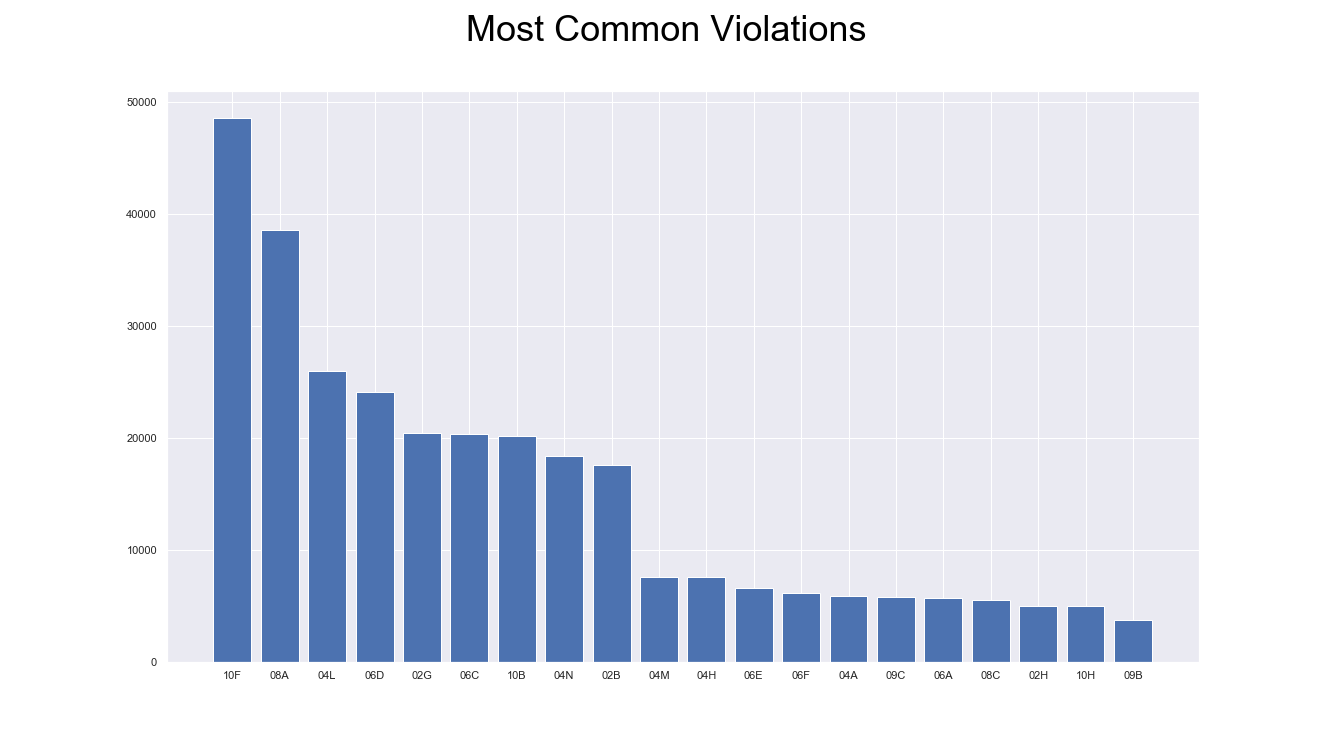
Another thing to note. Immediately following an inspection the inspector may offer a settlement for some violations in the form of a reduced penalty, provided you pay before your scheduled hearing day and don’t contest the violation. This is probably a good idea.

Worst case scenario is failing to appear, in person, to your hearing. This results in a default penalty on all cited violations. When a respondent is found in default, the penalty for each violation of the Health Code or the State Public Health Law must be doubled, except that in no case may the penalty imposed exceed $2000. This requirement of appearing in front of a hearing officer puts an additional drain especially on owners of small, independent FSEs. First, they may be unaware of such a rule but also they may not be able to take a block of time off during the middle of a weekday to go downtown for a hearing.

**The violations:**

The 9 most commonly cited violations account for over 50% of the total. The single most common itself is given nearly 15% of the time. Code 10F:

* Non-food contact surface improperly constructed. Unacceptable material used. Non-food contact surfaces or equipment improperly maintained and/or not properly sealed, raised, spaced or movable to allow accessibility for cleaning on all sides, above and underneath the unit.



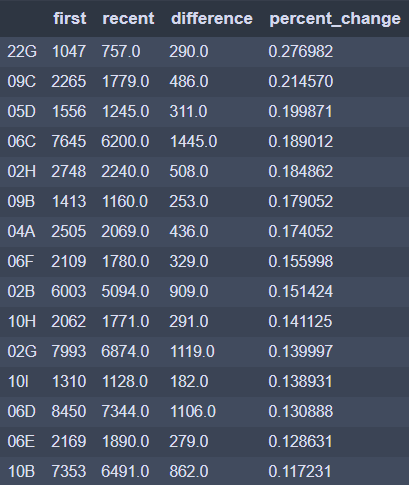
<https://www1.nyc.gov/assets/doh/downloads/pdf/rii/ri-violation-penalty.pdf>

Pretty nebulous, huh? It appears this is a catch all. In my personal experience this was given for a malfunction refrigerator. So, keeping all equipment in clean and working order is a likely way minimize your score.

Also of note in the top 9 are three mentions vermin and other pests. Make sure to have a reputable pest control vendor making regular passes through to catch any issues before an infestation takes hold.

Also, ice machine

It isn’t apparent in the dataset, but one can deduce which violations are overturned in the administrative hearing. According to the opendata page itself, the data consists of all sustained or not yet judicated violation for active restaurants. With that in mind I have been saving a fresh copy of the dataset almost daily since March. With this dynamic, historical capability I can see the first reporting of an unjudicated violation and compare that with a version from a later date. Here is a table of the most commonly overturned citations which have occurred over X times. Consider this if one of these happens to you.



**The chain effect:**