Ka Ieng Shery Cheong

June 20, 2017

Analysis of NYC Restaurant Inspections

Introduction

The data set containing inspection ratings for NYC restaurants collected between 2013 and 2017 reveal some very interesting insights about how factors like cuisine, borough, years of operation, and even seasonality influence the overall inspection score given. The lower the score, the better the grade, and the safer the restaurant is deemed to be. In this analysis, we will explore some of these factors and provide some recommendations to both consumers and inspectors on how to make the most of their visit.

Methodology

The bulk of the data analysis will be done via the pandas library in Python. Although the dataset provided is very normalized, the data is spread out over multiple files which requires us to first merge the files together using a series of joins. Next, some data wrangling is needed to create some useful variables such as a flag for whether the inspection visit occurred in a "warm" or "cool" month. We'd also like to convert granular data points like the inspection date to month so we can get a big picture view of trends over time.

After we create all the necessary variables, we'd like to aggregate the data by various factors such as month, cuisine, and borough. This is because the original data is at the inspection level, and we'd like to reduce the noise by looking at summary statistics like the average inspection score. This is done using the handy 'group by' function in pandas. The final output for this analysis will be based on the aggregated form of the data, and will be presented visually as tables and plots created using matplotlib.

Insights & Recommendations

Top 5 cuisine types with highest food safety

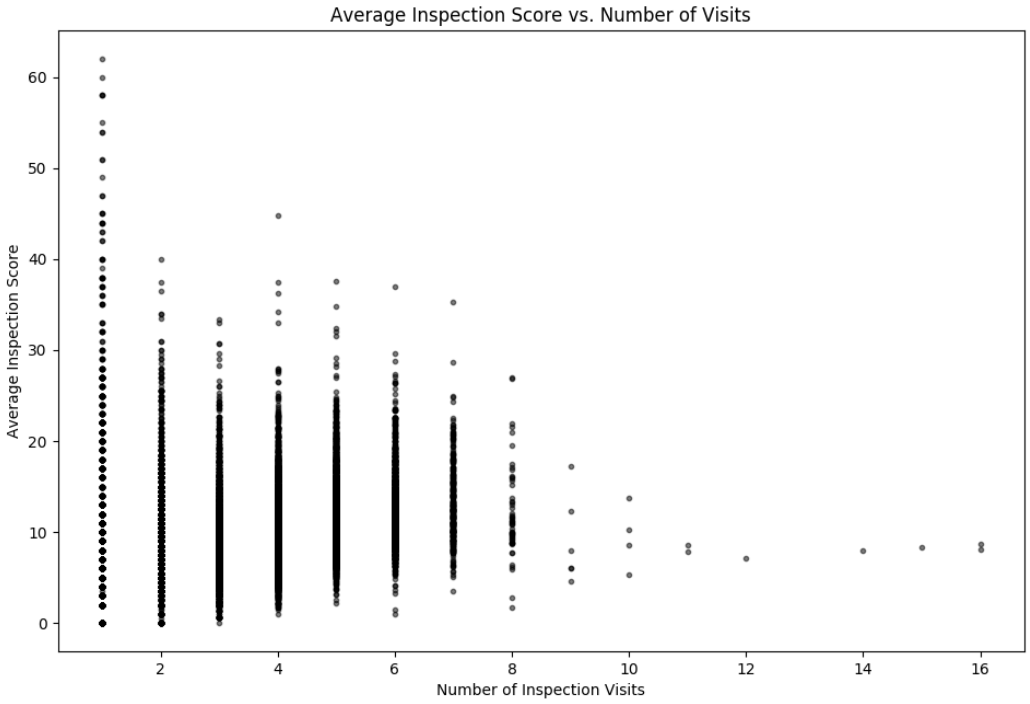
cuisine_description	grade	critical_flag	score	num_visits	num_restaurants
Nuts/Confectionary	A	0.479167	4.875	2.25	4
Soups	A	0.433333	4.9	2	5
Hotdogs/Pretzels	A	0.666667	7.13725	2.58824	17
Hotdogs	A	0.807692	7.28462	2.92308	26
Donuts	A	0.83916	7.8374	3.61209	513

Top 5 cuisine types with lowest food safety

cuisine_description	grade	critical_flag	score	num_visits	num_restaurants
Chilean	A	1.74762	15.4476	3	7
Polynesian	B	1.825	14.875	4.5	2
Creole	A	1.5903	14.2538	4	27
African	A	1.39619	12.8376	3.65714	70
Peruvian	A	1.46703	12.4767	3.89041	73

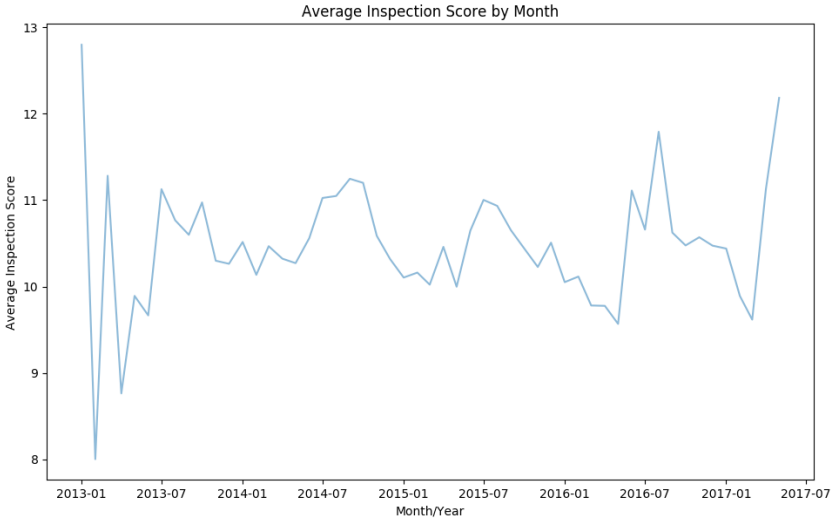
The cuisine types with generally low scores (and therefore, high safety ratings) are snack stores that sell food like candy, hotdogs, or donuts. Many of these tend to be chain stores, which might explain the more rigorous food safety protocols that are in place.

On the other hand, the cuisine types with higher scores (which translate to lower safety ratings) are ethnic eateries, which are more likely to be independently owned establishments. This is not to say, however, that consumers should avoid these restaurants, since 4 of 5 “worst” cuisine types still got an average grade of “A”.



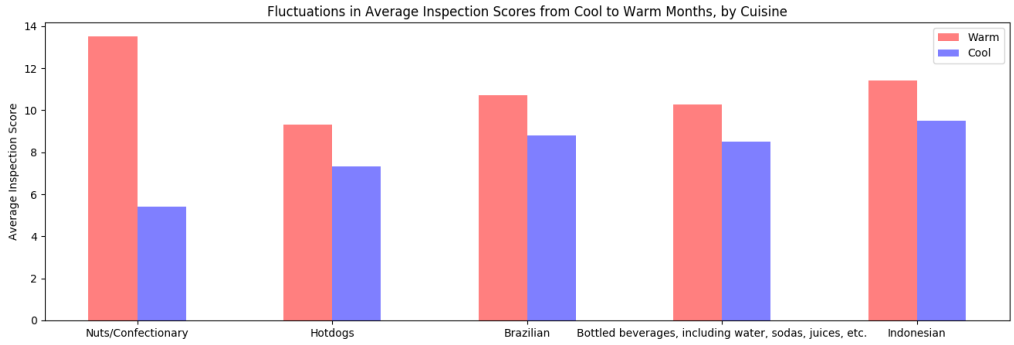
The next section we will explore is the correlation between the average inspection score and the total number of inspection visits for each restaurant in the data set. We see a larger variance in inspection scores for restaurants who have only been inspected once, but overall lower scores for restaurants who have had more inspections. This however, does not mean that scheduling more inspections lead to cleaner restaurants. If the inspection occurs only once or twice a year, this would mean that restaurants with a high number of visits have been around for a longer period, and improved their safety compliance over time.

<u>Recommendation for consumers:</u> More established restaurants are more likely to be safer than newer restaurants and will have less fluctuations in safety ratings.



The above graph shows the average inspection score plotted against the month in which the visit occurred. We can see that generally, the score spikes during the warmer months. This makes sense, as safety violations involving food temperature and vermin control will increase during the summer.

Recommendation for inspectors: Given the fluctuations in scores based on seasonality, it would not be fair for certain restaurants whose visits are always scheduled for the warmer months. Instead of making visits completely random, alternate visits between cool and warm months to ensure a more well-rounded score.



Finally, we'd like to delve deeper into how seasonality affects food inspection ratings. Are certain cuisines more susceptible to greater fluctuations once the weather warms up? It looks like for stores that sell candy, there is a drastic difference in ratings depending on when the visit occurs. This is interesting, given that earlier we saw that candy stores were actually in the top 5 cuisines with high safety ratings.

Recommendation for consumers: Avoid candy stores in the summer… unless you really, really like candy.