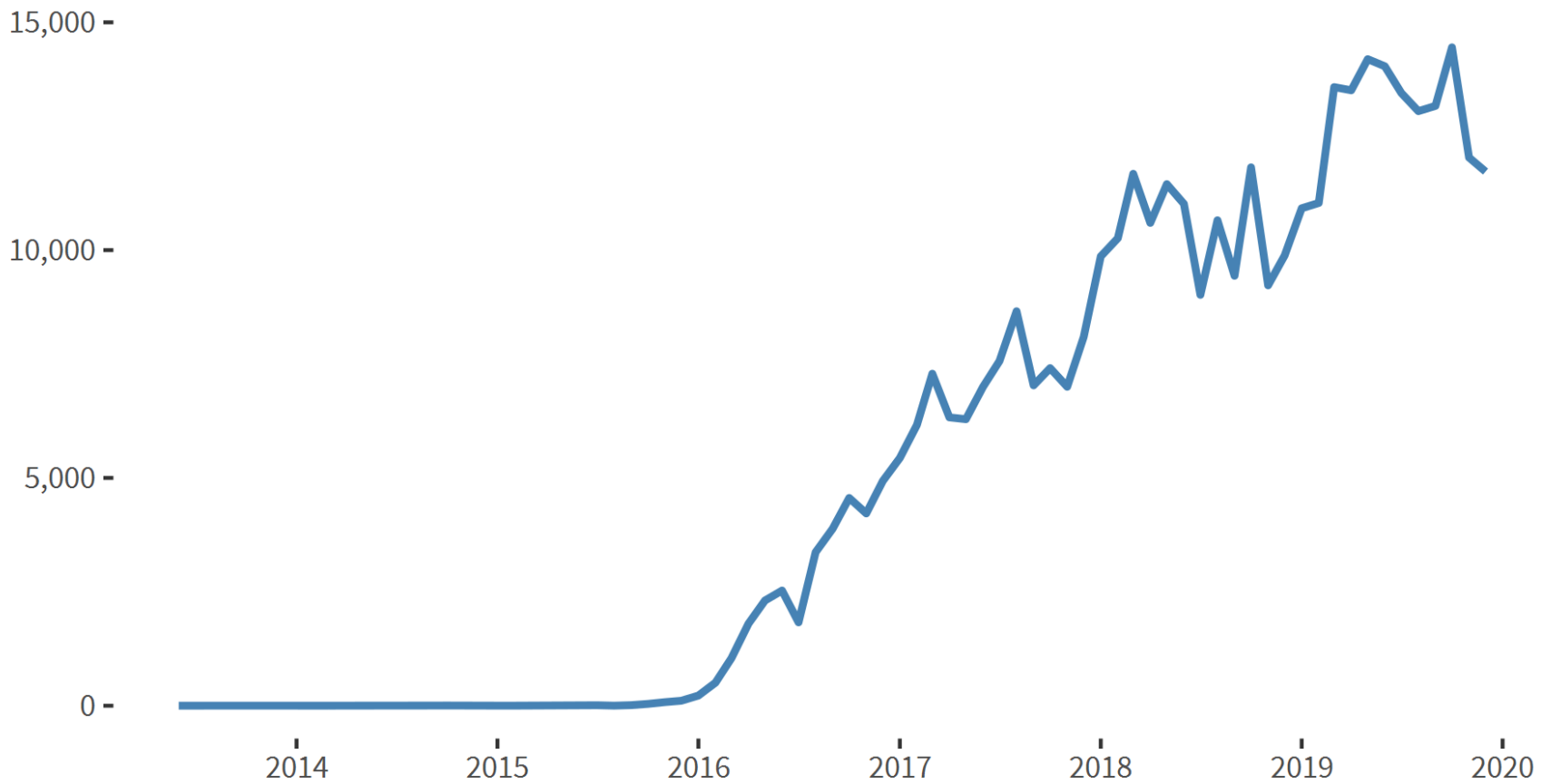


## Question 1

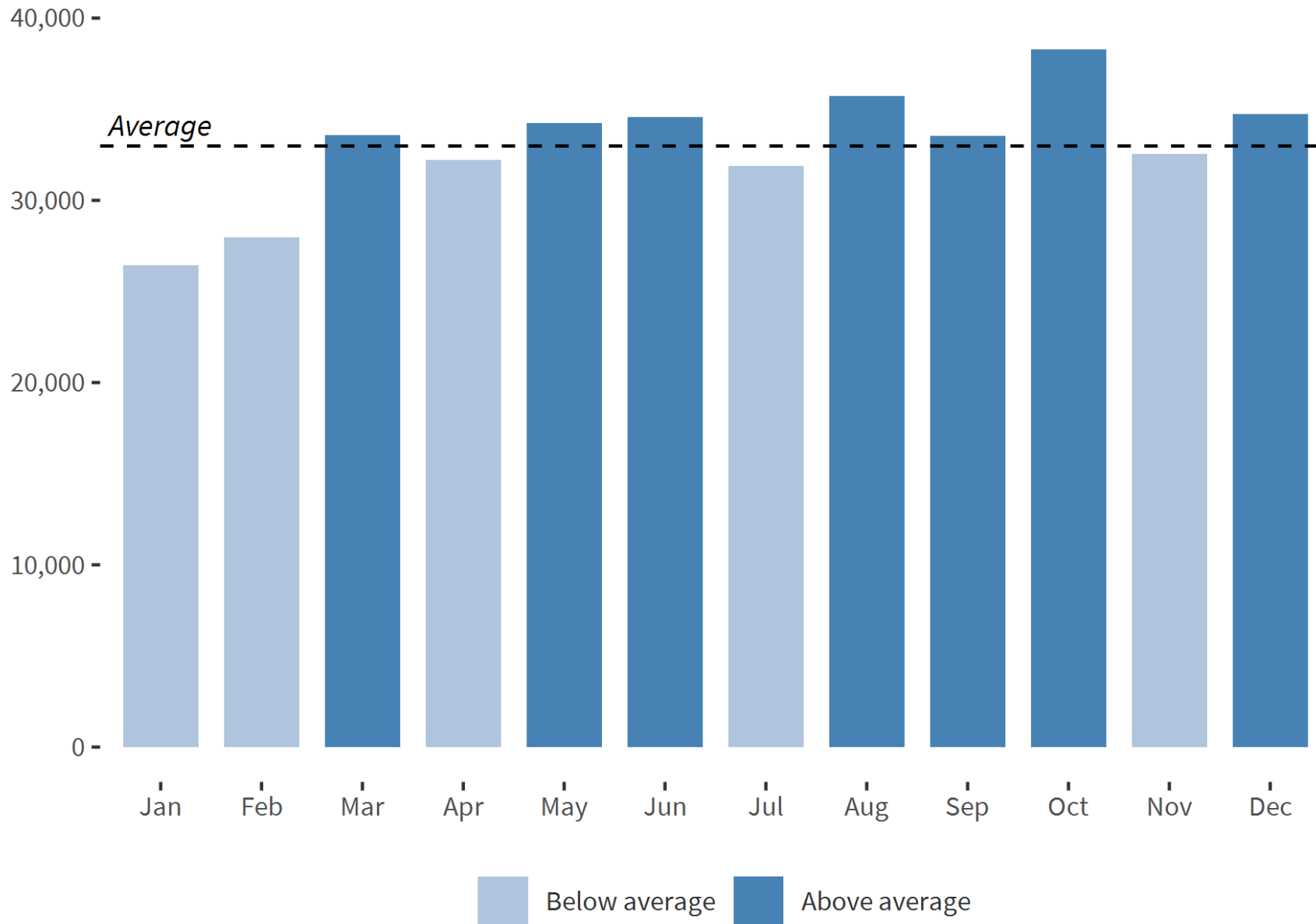
How does the number of inspections change over time (use month as the level of temporal granularity)? Does the number of inspections increase or decrease over time? Are there any peak times? Is there any seasonal effect (like inspections being more common during certain seasons or months)?

## Number of food inspections over time

The number of inspections increased sharply after 2016



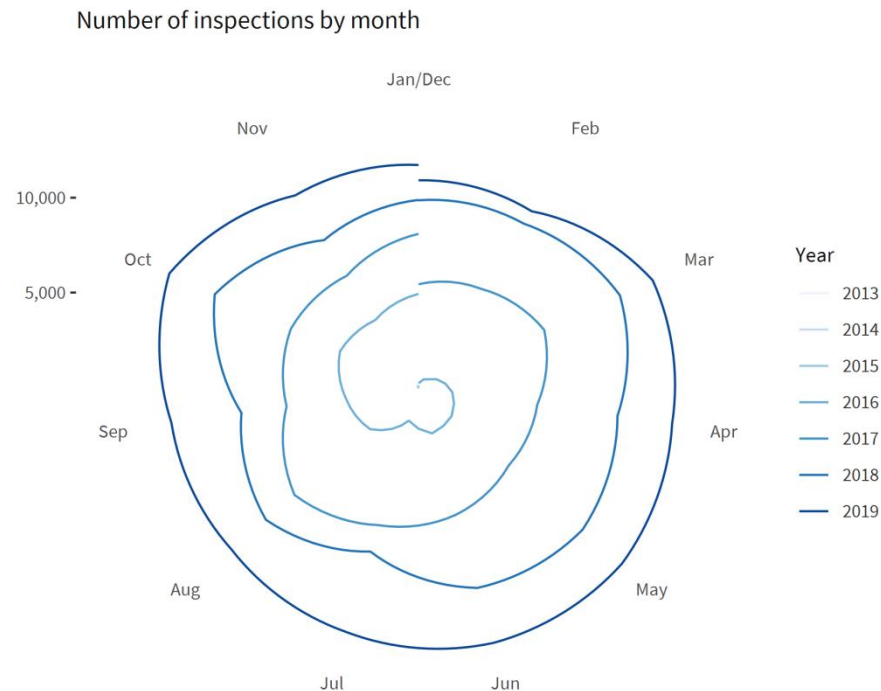
## Total number of inspections by month



# Comments

The line chart shows the number of inspections by month, in a continuous scale. A line chart is very well suited in this case, since it allows us to perceive patterns in time.

To answer the question about seasonal effects, I first transformed the line chart to polar coordinates, since it would probably show any evident recurring pattern along the years:



# Comments

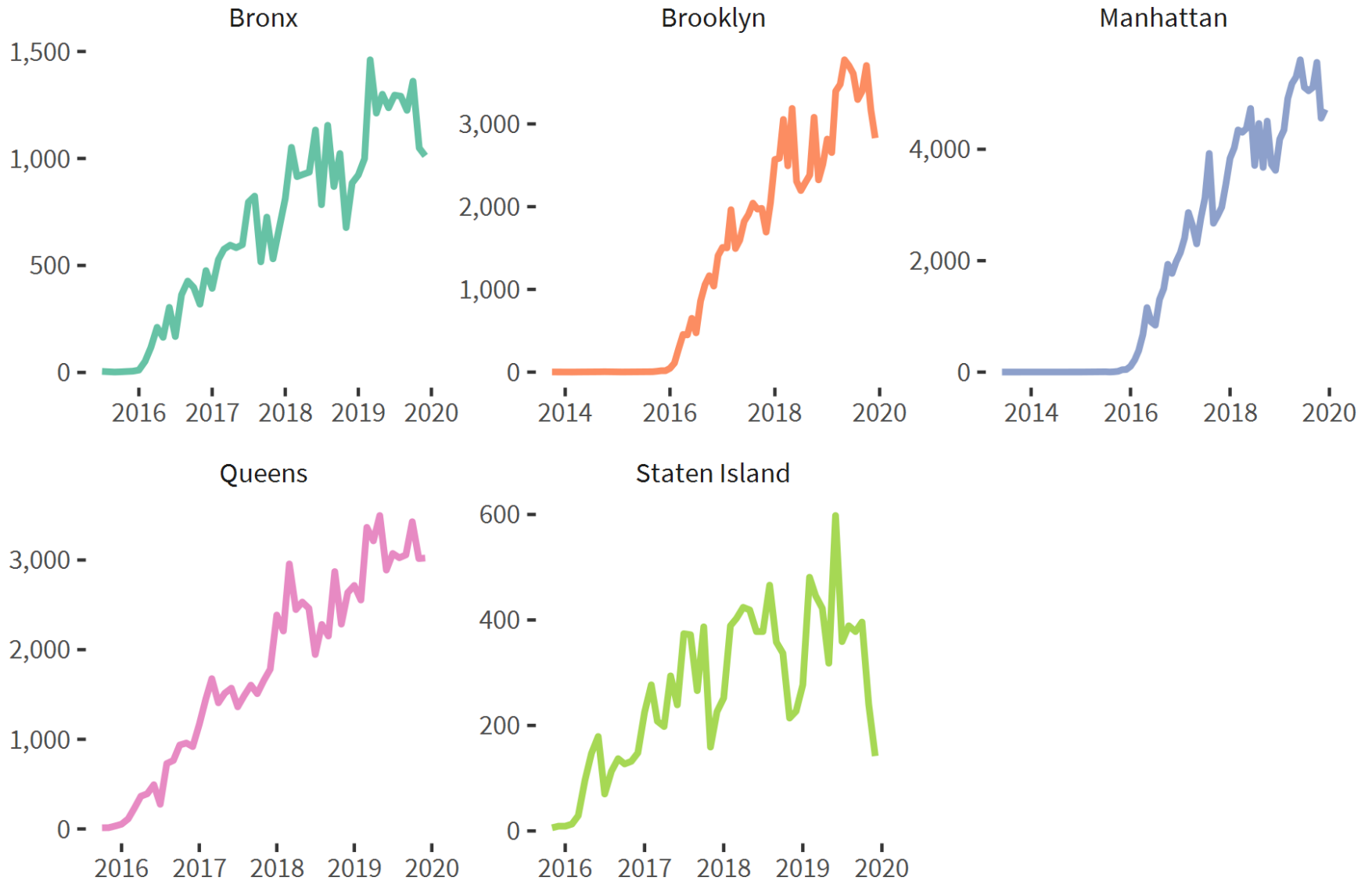
But there was no distinctive pattern that would “jump to the eye”, so instead I chose to design a more conventional bar chart, showing the total number of inspections by month, regardless of the year. Additionally, I included a line indicating the monthly average, highlighting the months above this average.

## Question 2

Is there any difference in how the number of inspections changes over time in the 5 different boroughs of New York City?

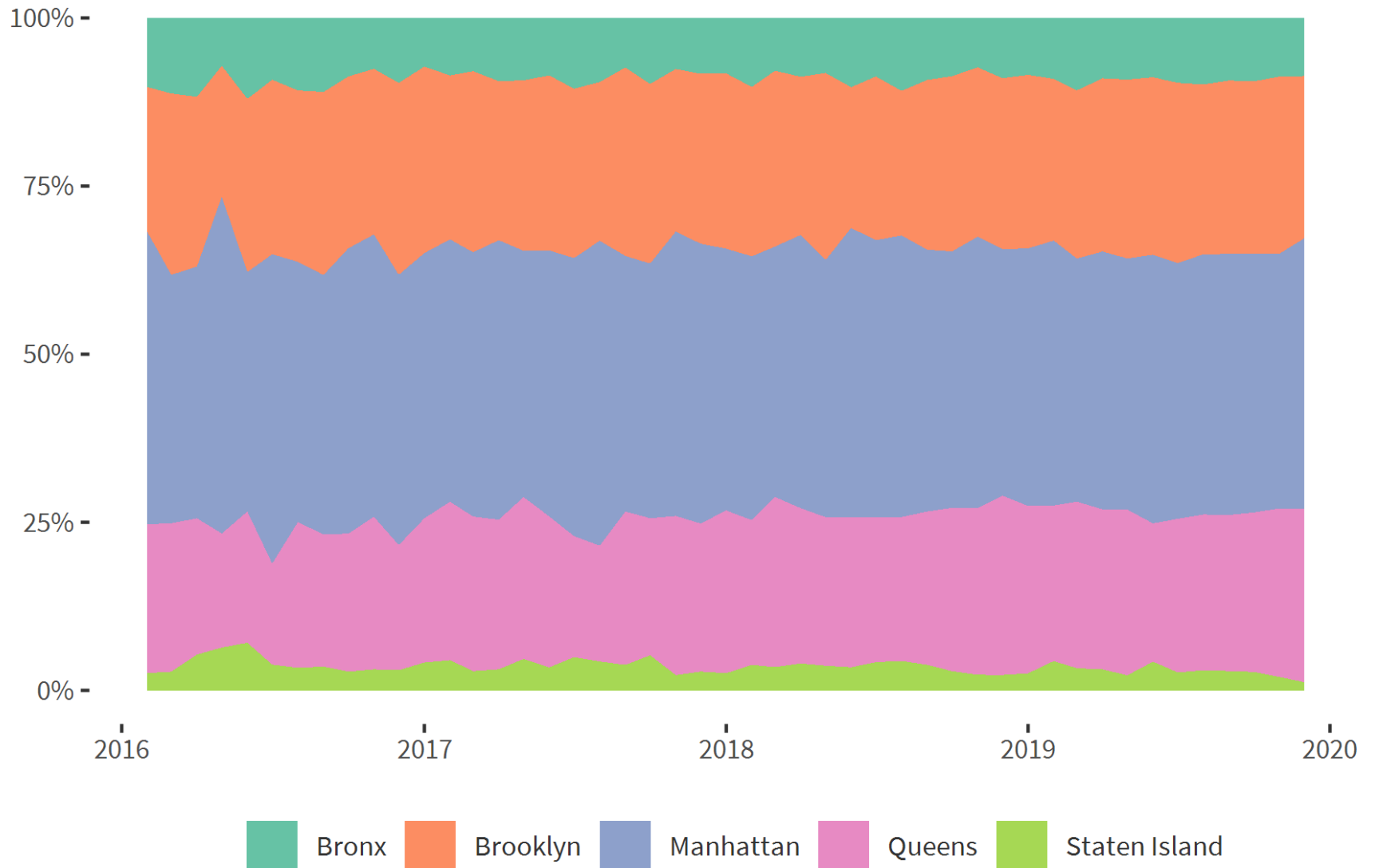
# Number of inspections over the years

By Borough



# Distribution of the number of inspections by Borough

From 2016 onwards

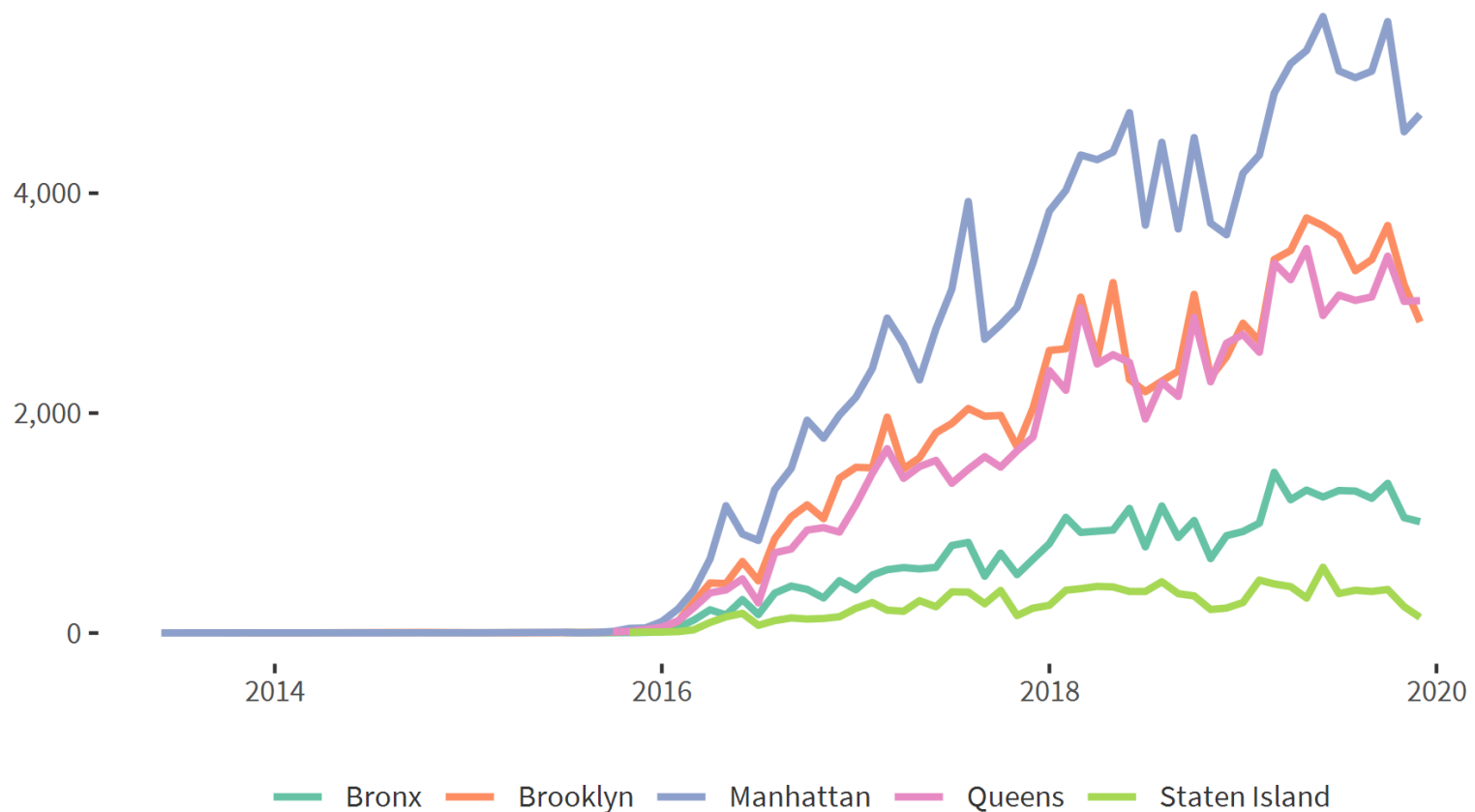




# Comments

Here, to show the patterns for each borough, the immediate idea would be to plot a line chart with multiple lines representing the boroughs :

Number of inspections over the years  
By Borough



# Comments

But I actually in the end chose to use a faceted version of the chart, with free scales, to show that the pattern is very similar for every borough.

To demonstrate this further, I also plotted the distribution of the number of inspections by borough, to show that the percentage of inspections at each borough is more or less stable over the years. I disregarded the years prior to 2016, because the number of inspections in general clearly took off on this year (see chart on the previous slide).

## Question 3

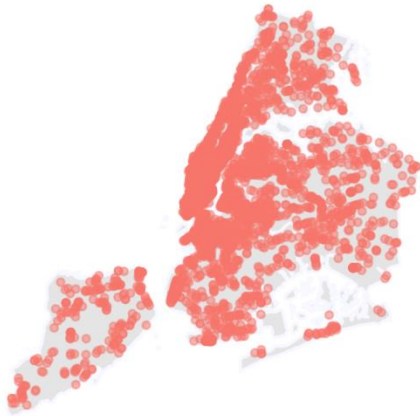
How are cuisines types distributed across the New York area?

Are there geographical areas where certain cuisines tend to concentrate (that is are there any areas where certain cuisines are more prevalent than others)? NOTE: focus only on the top 5 most frequent “Cuisine Description” categories.

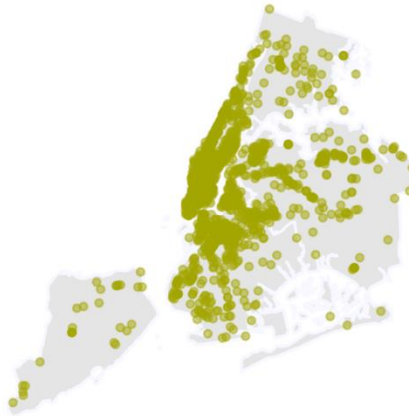
# Inspections by cuisine in NYC area

Only top 5 cuisines shown

American



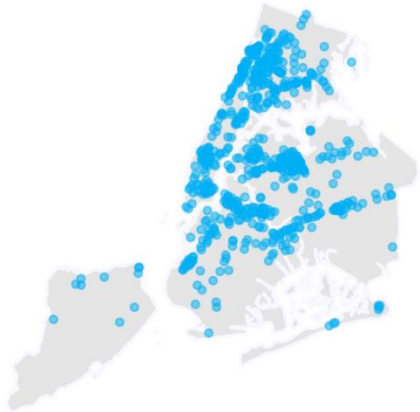
Café/Coffee/Tea



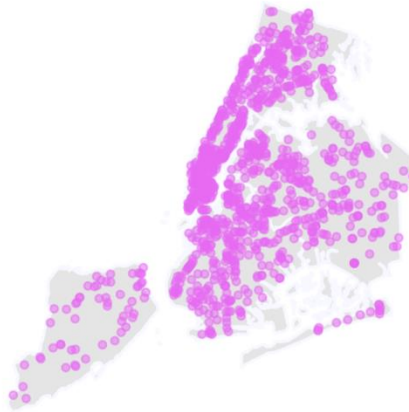
Chinese



Latin (Cuban, Dominican, Puerto Rican, South & Central American)



Pizza



# Comments

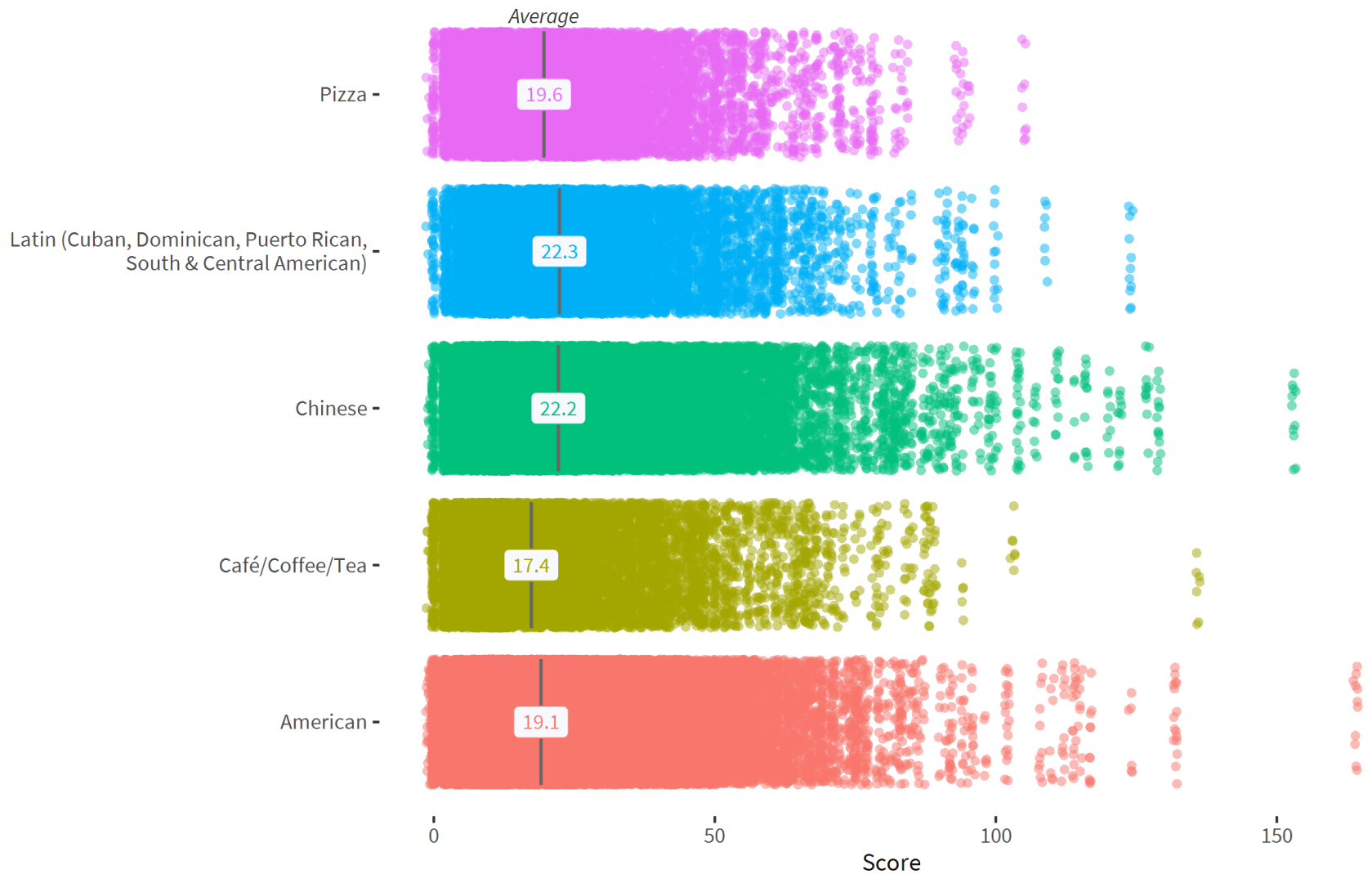
Again I chose to use a faceted version of the chart, to avoid cluttering.

## Question 4

How does the average score compare across different cuisine types? Are there cuisines that tend to have consistently lower/higher average scores compared to the others? NOTE: focus only on the top 5 most frequent “Cuisine Description” categories.

## Distribution of the Inspections's scores

Among the top 5 cuisines



# Comments

Instead of simply plotting the average scores by cuisines, I chose to also show the full distribution of the scores, since averages are sensitive to extreme values.



## Question 5

Is there a relationship between cuisine type and violation? For instance, do some cuisine types tend to have more of some type of violations than other cuisine types?

# Most frequent cuisine type for each violation

Considering the top 5 cuisines and the top 10 violations. Horizontal values (by violation type) add up to 100%.

	American	Café/Coffee/Tea	Chinese	Latin	Pizza
Raw, cooked or prepared food is adulterated, contaminated, cross-contaminated, or not discarded in accordance with HACCP plan.	61.9%	10.6%	13.4%	8.5%	5.7%
Plumbing not properly installed or maintained; anti-siphonage or backflow prevention device not provided where required; equipment or floor not properly drained; sewage disposal system in disrepair or not functioning properly.	53.1%	11.8%	17.0%	9.2%	8.9%
Non-food contact surface improperly constructed. Unacceptable material used. Non-food contact surface 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.	46.3%	13.0%	22.9%	7.9%	9.8%
Hot food item not held at or above 140° F.	33.7%	3.4%	32.2%	14.3%	16.4%
Food not protected from potential source of contamination during storage, preparation, transportation, display or service.	43.6%	10.9%	29.2%	8.6%	7.7%
Food contact surface not properly washed, rinsed and sanitized after each use and following any activity when contamination may have occurred.	56.3%	13.5%	15.4%	7.6%	7.3%
Filth flies or food/refuse/sewage-associated (FRSA) flies present in facility's food and/or non-food areas. Filth flies include house flies, little house flies, blow flies, bottle flies and flesh flies. Food/refuse/sewage-associated flies include fruit flies, drain flies and Phorid flies.	55.1%	10.5%	15.0%	11.1%	8.2%
Facility not vermin proof. Harborage or conditions conducive to attracting vermin to the premises and/or allowing vermin to exist.	45.9%	9.7%	23.7%	10.2%	10.5%
Evidence of mice or live mice present in facility's food and/or non-food areas.	41.9%	9.0%	27.4%	9.3%	12.4%
Cold food item held above 41° F (smoked fish and reduced oxygen packaged foods above 38 °F) except during necessary preparation.	45.7%	7.8%	27.5%	9.4%	9.6%

# Most frequent violations for each cuisine type

Considering the top 5 cuisines and the top 10 violations. Vertical values (by cuisine type) add up to 100%.

	American	Café/Coffee/Tea	Chinese	Latin	Pizza
Raw, cooked or prepared food is adulterated, contaminated, cross-contaminated, or not discarded in accordance with HACCP plan.	4.4%	3.4%	2.0%	3.0%	1.9%
Plumbing not properly installed or maintained; anti-siphonage or backflow prevention device not provided where required; equipment or floor not properly drained; sewage disposal system in disrepair or not functioning properly.	9.1%	9.2%	6.1%	8.0%	7.4%
Non-food contact surface improperly constructed. Unacceptable material used. Non-food contact surface 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.	20.2%	25.7%	20.7%	17.5%	20.6%
Hot food item not held at or above 140° F.	5.0%	2.3%	10.0%	10.8%	11.8%
Food not protected from potential source of contamination during storage, preparation, transportation, display or service.	8.0%	9.0%	11.1%	8.0%	6.8%
Food contact surface not properly washed, rinsed and sanitized after each use and following any activity when contamination may have occurred.	12.4%	13.5%	7.0%	8.4%	7.7%
Filth flies or food/refuse/sewage-associated (FRSA) flies present in facility's food and/or non-food areas. Filth flies include house flies, little house flies, blow flies, bottle flies and flesh flies. Food/refuse/sewage-associated flies include fruit flies, drain flies and Phorid flies.	8.7%	7.6%	4.9%	8.9%	6.3%
Facility not vermin proof. Harborage or conditions conducive to attracting vermin to the premises and/or allowing vermin to exist.	15.2%	14.5%	16.2%	17.1%	16.7%
Evidence of mice or live mice present in facility's food and/or non-food areas.	9.0%	8.8%	12.2%	10.1%	12.8%
Cold food item held above 41° F (smoked fish and reduced oxygen packaged foods above 38 °F) except during necessary preparation.	8.0%	6.2%	9.9%	8.3%	8.0%

# Comments

Here, I first plotted, for each infraction, the cuisine types that incurred in more infractions. But since “American” cuisine is the most prevalent cuisine in the dataset, the plot would be somewhat unbiased.

To fix that, I decided to plot a chart showing the most frequent violation for each cuisine type. The difference is subtle, but the resulting plot is very different from the first one (and more interesting).