[Use this template for your final project description. Then produce a PDF version and submit it]

[Please use proper English and be as precise as possible]

NYC Food Inspection

*Visualizing NYC’s restaurants’ sanitation score based on location and cuisine type*

Sijie Liu sl5039 [sl5039@nyu.edu](mailto:sl5039@nyu.edu)

Jiming Ye jy1769 [jy1769@nyu.edu](mailto:jy1769@nyu.edu)

Yuchi Du yd828 [yd828@nyu.edu](mailto:yd828@nyu.edu)

Liuming Zhang [lz1074@nyu.edu](mailto:lz1074@nyu.edu)

Project page (on Github): … [add url]

Video: … [add url]

Working demo: … [add url]

**What is the problem you want to solve and who has this problem?**

This project is intended to inform customers about the NYC’s restaurants sanitation inspection results. It will help customers decide where to eat based on restaurant's grade and violation severity.

**What are the driving analytical questions you want to be able to answer with your visualization?**

* Question 1: Find restaurants with different violation grades based on their address.
* Question 2: For every restaurant, find the grade change tendency through different times.
* Question 3: For every restaurant, we can see its cuisine description
* Question 4: For every restaurant, we can find out the violation description cited in the different areas
* Question 5: For each type of restaurants, we can find out the average score of the certain type of restaurants.
* Question 6: For each borough in NYC, find out the average score within its area.

**What does your data look like? Where does it come from? What real-world phenomena does it capture?**

​​

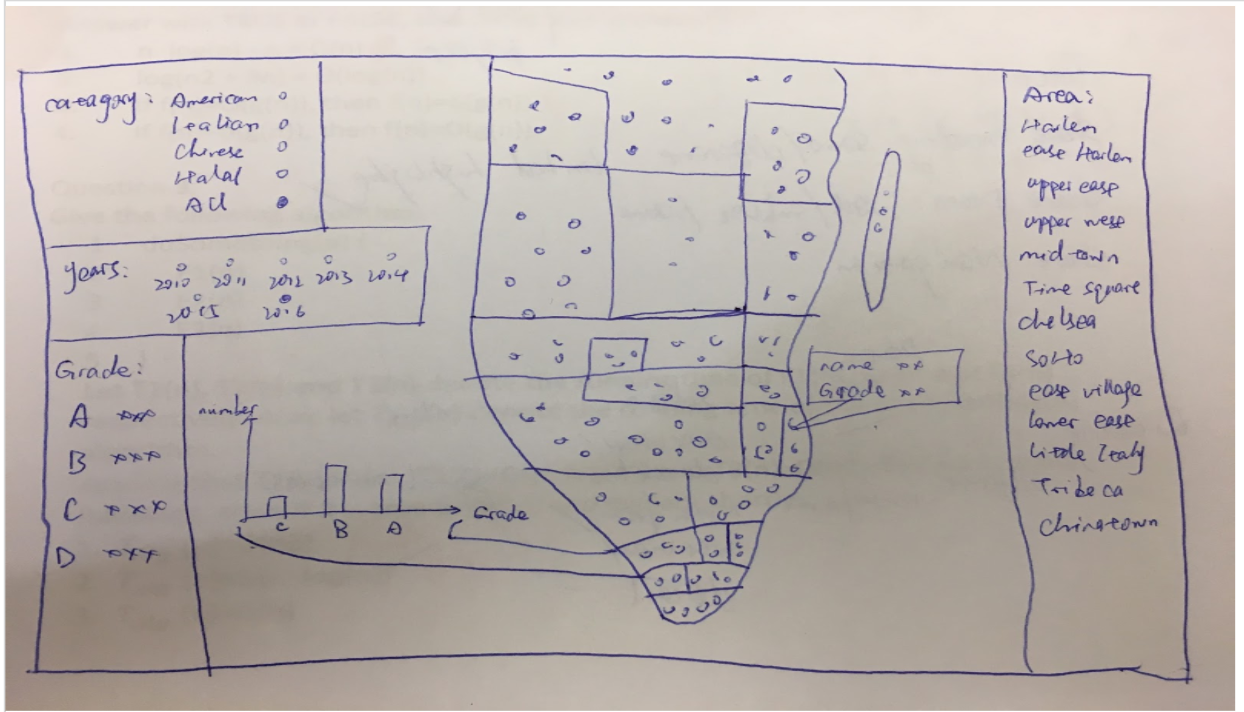
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Attribute Name | Attribute Type | Meaning | Values | Derived? (if yes explain how) |
| camis | categorical | Restaurant ID | number | No. |
| dba | categorical | Restaurant name | string | No. |
| boro | categorical | address | string | No. |
| building | categorical | address | string | No. |
| street | categorical | address | string | No. |
| zipcode | categorical | address | number | No. |
| Cuisine description | categorical | Food types | string | No. |
| Violation code | categorical | Violation types | number | No. |
| Violation description | categorical | Descriptions for different violation codes | string | No. |
| Critical flag | categorical | Whether the violation is critical. | bool | No. |
| score | quantitative | Hygiene inspection score. | number | No. |
| score\_avg | quantitative | Sum(score)/amount | number | Yes. |
| grade | ordinal | Restaurant inspection grade | char | No. |
| Grade date | ordinal | Date restaurants get grade. | date | Yes. Extract year and month from original grade date. |

**What have others done to solve this or related problems?**

Describe related works and explain how they are related to your work.

**Design Iterations**

Attempt 1:

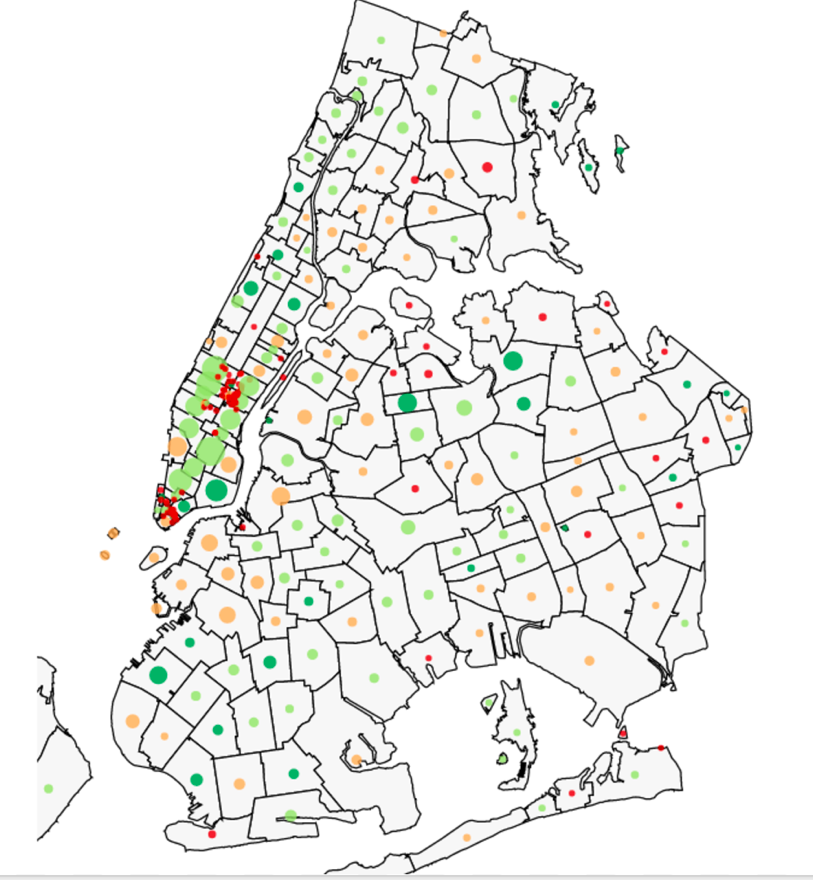


Our first attempt is to show every restaurants as a single dot in the map, and show all these calculated results in the blank area. Using different color to represent different grades, also user can filter the result by different cuisine types and by different year. Besides, there will be statistic bar chart in the main page, which shows the number of restaurants with different grades.

The obstacle:

* There are too many restaurants, can not put each restaurant as a single dot
* Too many information showed in the main page
* Map choice

Attempt 2:



We draw this map using D3 js, the different dot with vary colors means the number of restaurants in this borough and the average sanitation scores of this area.

In this case, every borough will have one dot only, the size represents the restaurant number, which is clear to see, the dark green color means the average grade is A, light green is B, red means C and lower.

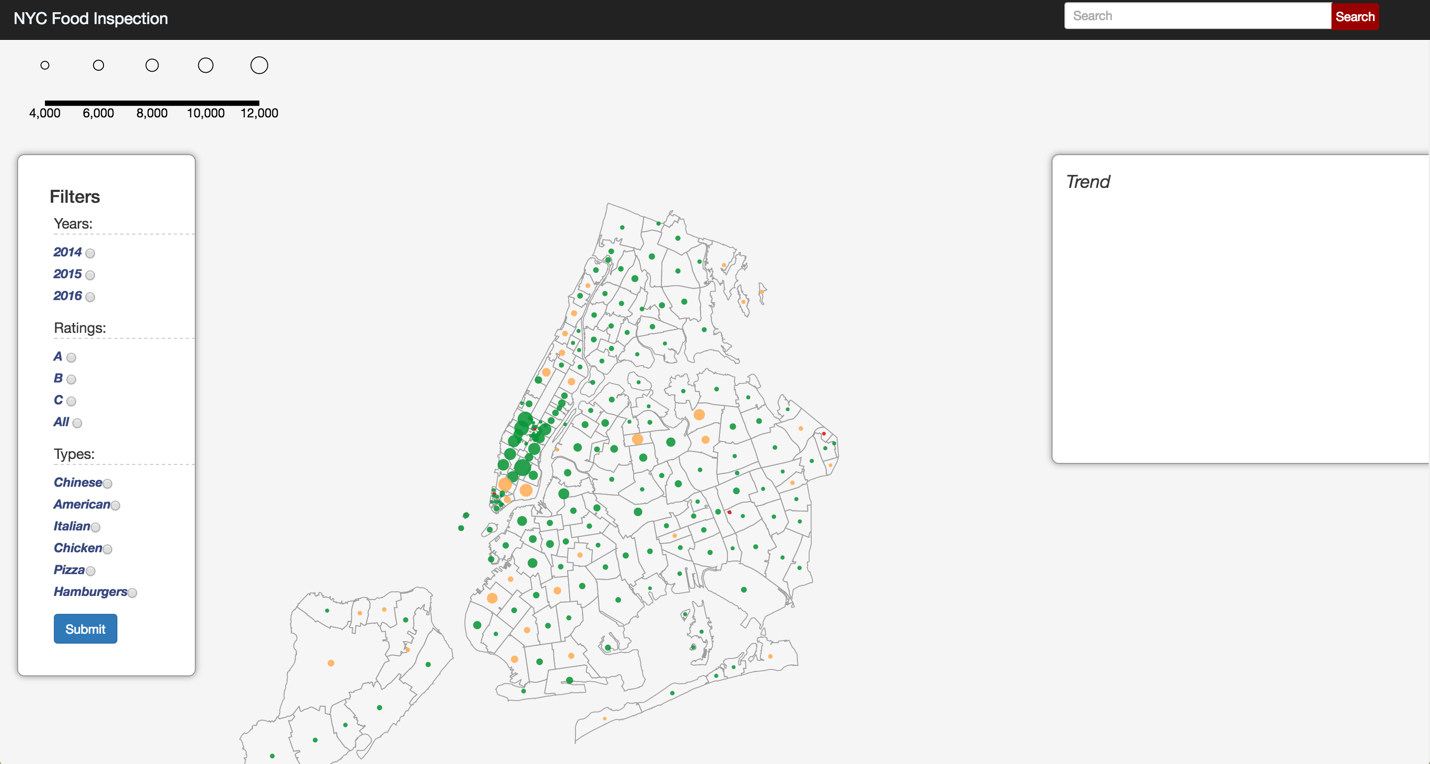
Obstacle:

* No filter function aside the map
* No specific number of restaurants and composition of each dot
* No search function to search a specific restaurant
* Some built-in APIs doesn’t work as expected

Future works:

* Add a pop-out pie chart for each dot
* Add a searching bar
* Add year, grade, cuisine type filters

**Final Visualization**

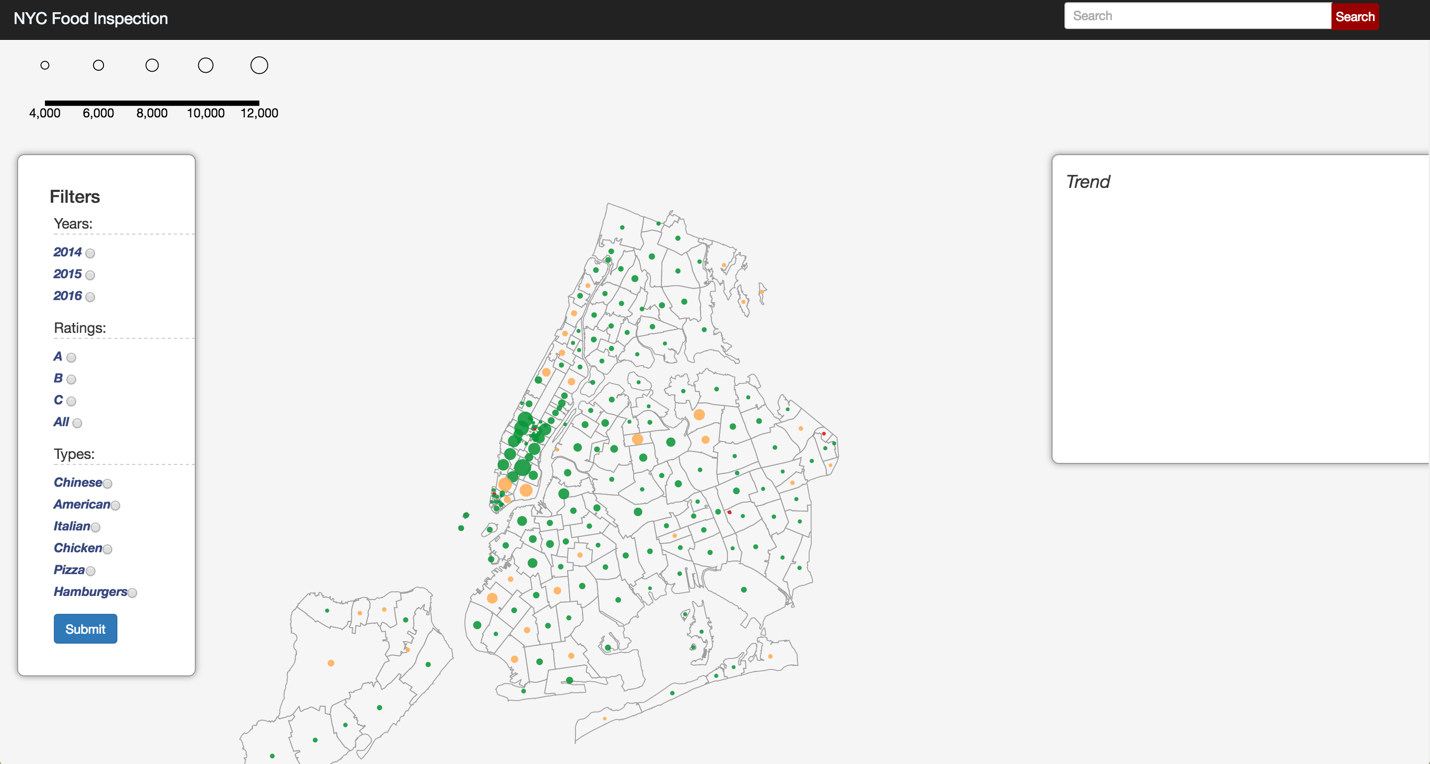


The size of spots on the maps represent the quantity of restaurants in this specific borough.

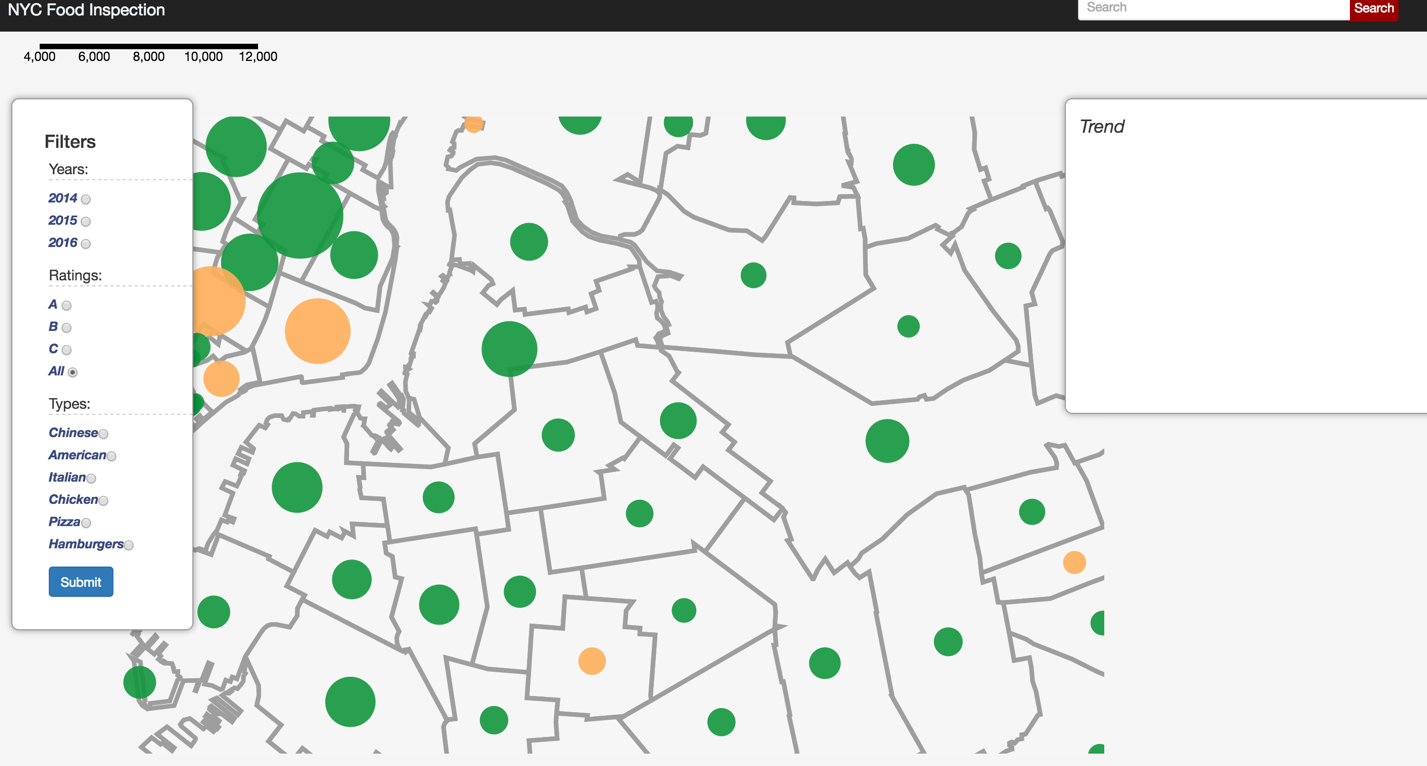
The color represents the average sanitation grades of

Insert images of your final visualization, describe how to read it

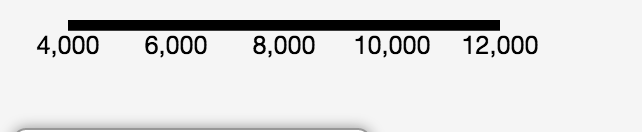
**Findings**



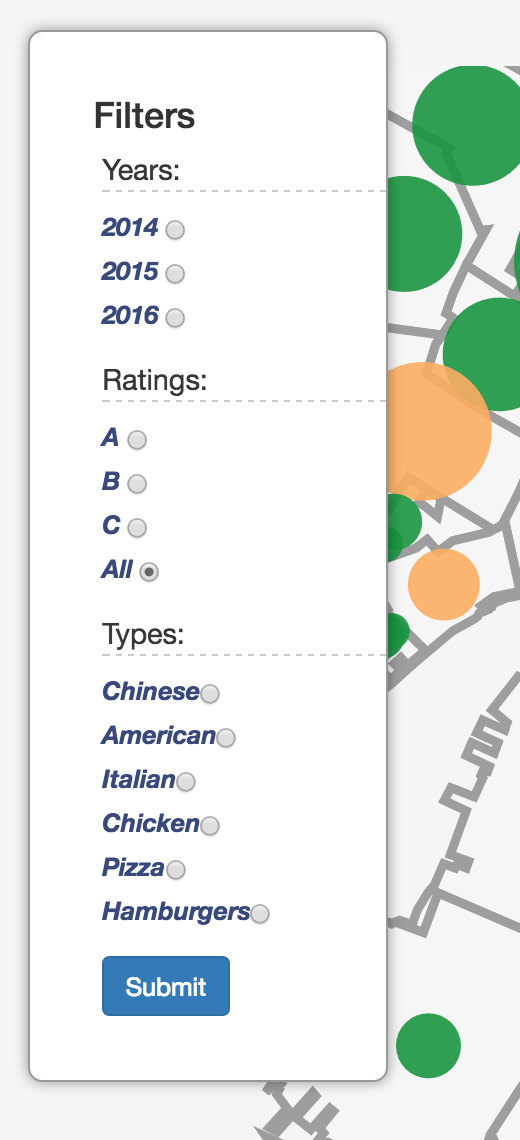
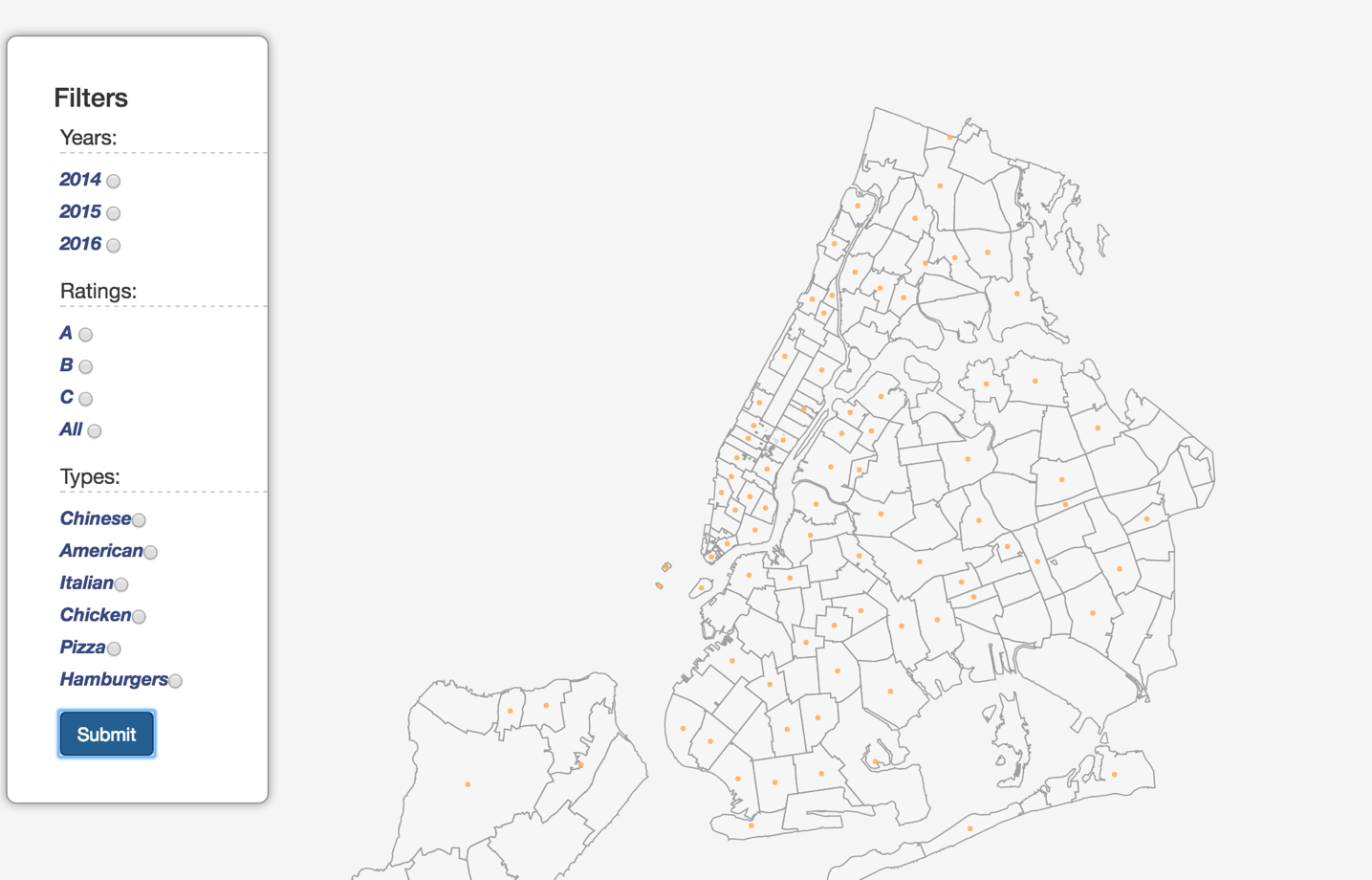
* From a overall view of our project, we split NYC into different boroughs by zipcode, extract restaurant data and its grade, calculate the average grade and show it on the map as a single spot.
* We use four colors to represent the level of grade, and the area of the spot stands of the average grade.



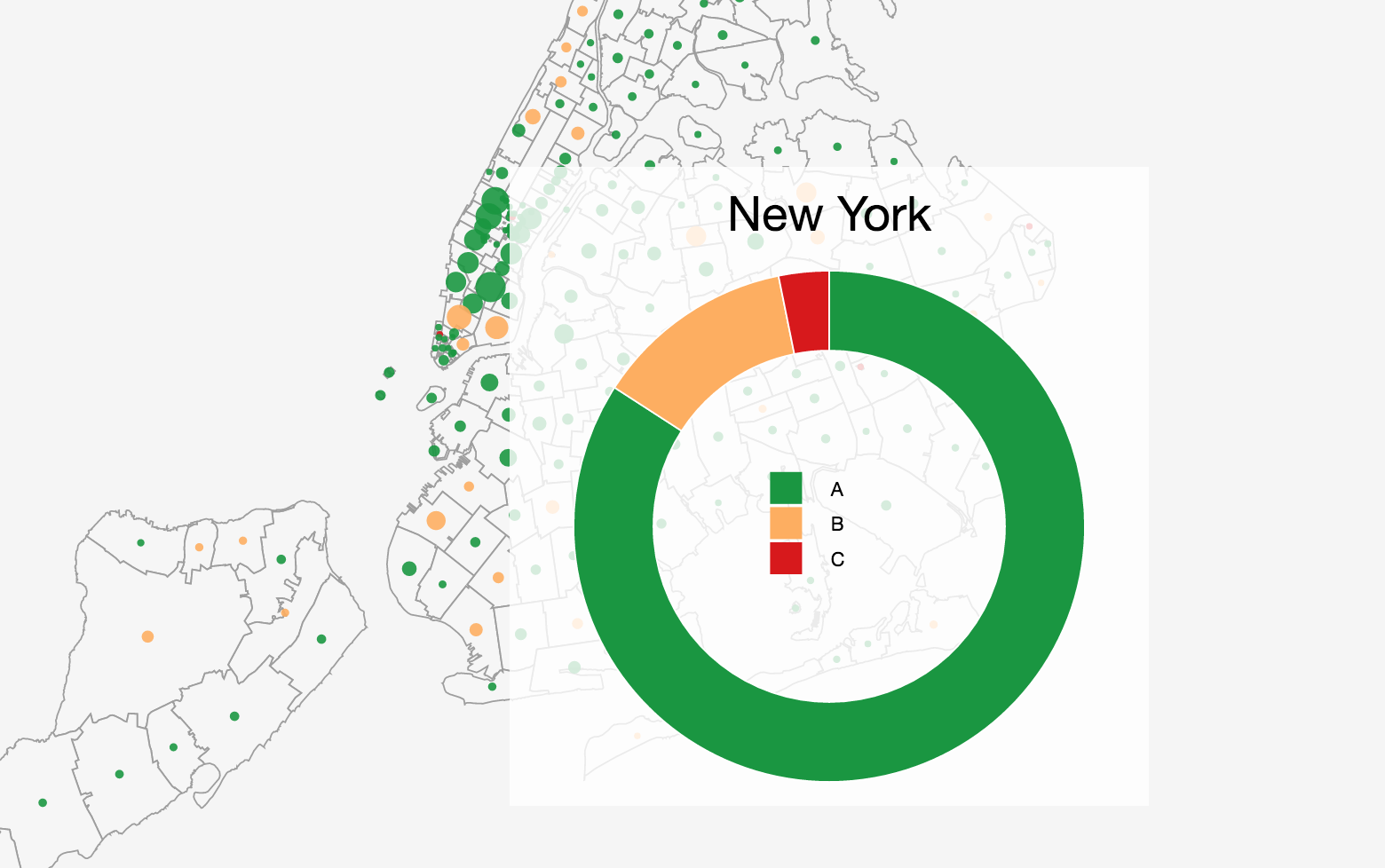
* As user click the empty area, it will automatically zoom in, it shows the detailed map and spot.



* The legend, which represents the relationship between the size of circle and the quantity of restaurants, at the top left of the page.

* In the left of the page, we add a filter bar to support filter functions which can help user to get results for different conditions. It is a combination filter bar. Users can select multiple conditions and apply them to filter data.



* When user click one spot, it will pop up a pie chart. It shows the percentage of different grades in this boroughs.



* We add a search function with autocomplete to support the user to find one specific restaurant situation. When users search the restaurant name, it will be showed on the map and its grade’s trend will be represented at the trend area in line chart.

[NOTE: This section is very important for your final submission! If you have doubts on how to do it right contact me as soon as possible!]

Describe the results of your data analysis performed with your application. What interesting information did you extract? What are your major discoveries? What did you learn about the data? What interesting data stories can you show using your developed tool?

Please add as many screenshots as needed and describe how to read them and what interesting information can be extracted from them.

I suggest you to organize this as progression of questions and screenshots. Start with a question and show it can be answered with your tool. Start with one that has an overview, describe what can be observed there, then how this leads to a new question, then how this leads to a new question, then a new screenshot showing how this question can be answered and so on ...

**Limitations and Future Works**

What are the major limitations of your project at this stage (it’s fine to talk openly about what does not work well yet and what you have not been able to realize)? If you had more time to develop your application further, what would be the next steps?

Currently, this work is not perfect because it cannot show more detail information for one specific cuisine type and cannot show some complex information. It only offers some general information such as average grades for different boroughs. So if we have more time, we want to do more calculations on data to find relationship between cuisine type, district, and grade. We may find more valuable information.