NYU | Center for Urban Science + Progress

CUSP-GX-6006.001: Data Visualization

SPRING 2018

Homework 2

Due Date: 5:30PM on Mar 22, 2018

(UPDATED) INSTRUCTIONS: In this homework, we will review plotting principles, and building an interactive spatial visualization to explore restaurant cuisines in New York City. For each task, you must turn in a single URL that can be accessible from the Internet to show your results. This could be:

- A link to your working JS Bin (you must have our JS Bin and GitHub accounts linked up to do this)
- A publicly accessible web page showing your visualization hosted on, for example, GitHub Pages, bl.ocks.org, or a combination of GitHub Gist and rawgit.com.

TASK 1 (10 points)

Please identify a plot or visualization that are available through online media, or in printed materials, where:

1. It makes use of any of the data sets available on NYC Open Data, ***and***
2. You believe that the visualization can be improved based on **two or more** of the plotting principles that we have discussed in class (i.e. the 4 principles for improving the vision, and the 5 principles for improving the understanding).

After identifying such a visualization and its potentials for improvement, you are asked to build an improved version that addressed the using the same data and D3.js. In this task you are asked to turn in:

1. A document (could be a PDF or a Word document) that describes the visualization including the data source, the purpose of the visualization, its problems (of which plotting principles they violate), and how they can be improved.
2. A link to your D3.js version of the visualization as described in instructions in the beginning.

TASK 2 (10 points)

For this task, we would like to explore the different cuisines that restaurants in New York City have to offer by zip code. You are asked to build an interactive visualization that:

- Display the top 25 popular cuisines in NYC (ranked by the number of restaurants of those cuisines). We assume that the more restaurants there are for a cuisine, the more popular it is.
- Allow users to select any of the 25 top cuisine, and display the distribution of restaurants of that cuisine across zip codes in NYC.

A target visualization is provided in below. In Figure 1, a user can browse through the top cuisines ranked by the number of restaurants on the left. The image shows that when the user hovers the mouse over the American cuisine, the map on the right displays the distribution of the cuisine across all zip codes. On the top left corner of the map, there is also a color scale with an appropriate title indicating the scale for each zip code. In this case, the mid and lower part of Manhattan have the most American restaurants, with some zip codes consisting of almost 3,000 restaurants. In Figure 2, when the user moves the mouse over to Chinese cuisine using the bar graph on the left, the map on the right is updated accordingly, showing the distribution of Chinese restaurants throughout NYC. They are mostly in China Town in Manhattan, Queens, and Brooklyn with close to 2,000 restaurants in some zip codes.

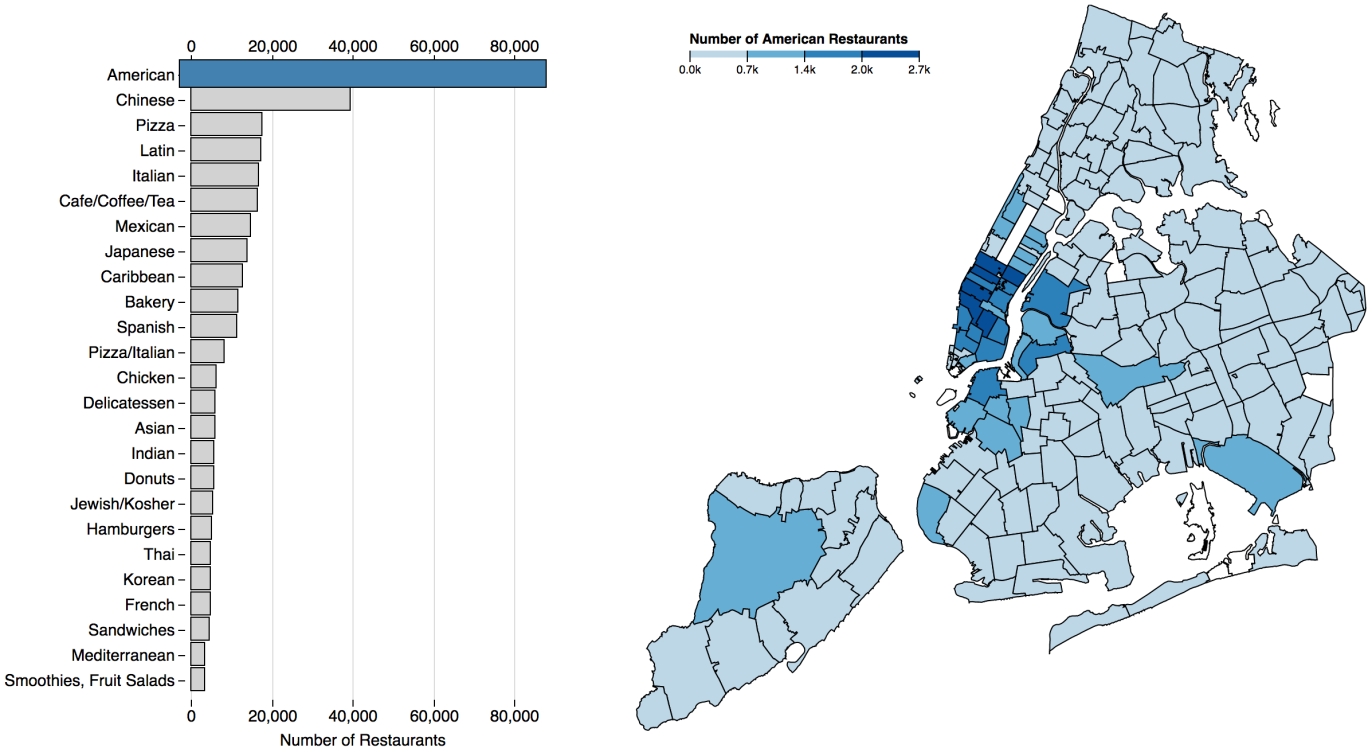

Figure 1. The distribution of American restaurants in NYC.

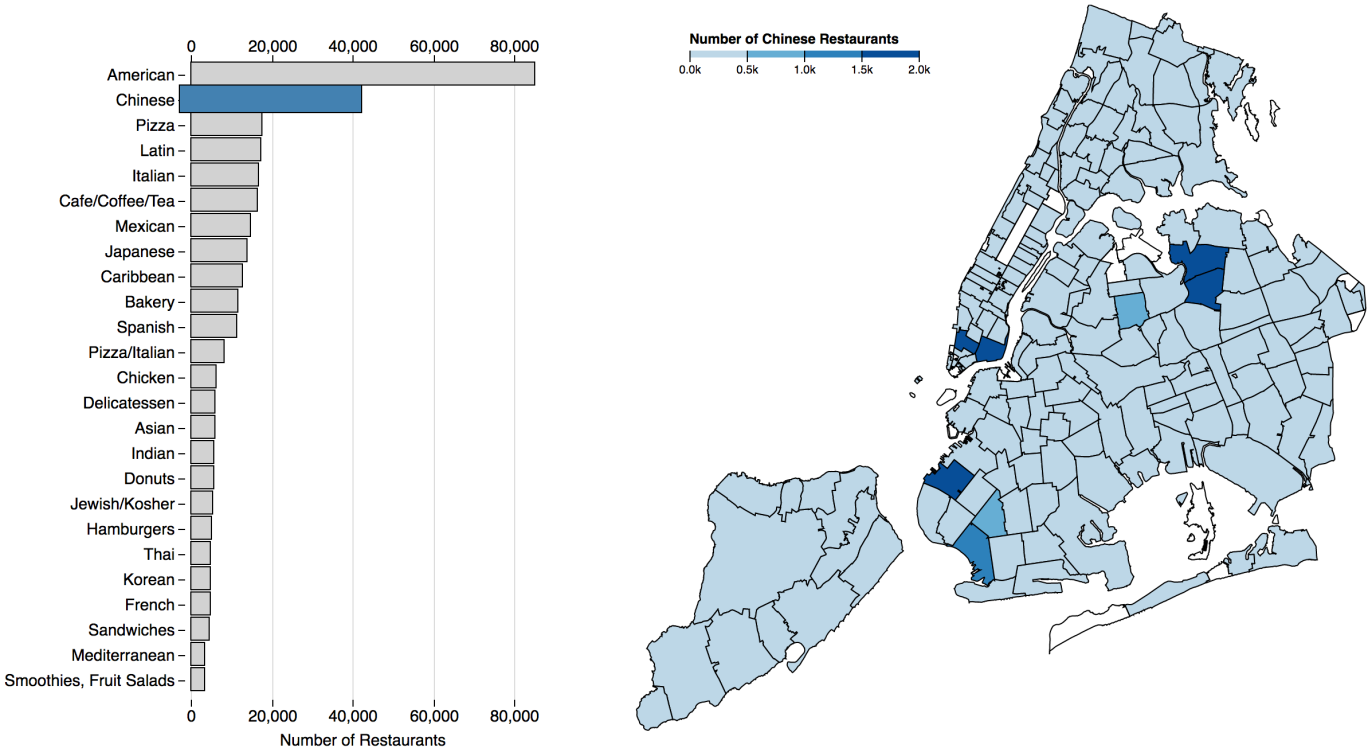

Figure 2. The distribution of Chinese restaurants in NYC.

We only need to display 5 different color ranges for the map (the four colors shown on the scale, and a transparent one, aka. "none"). All zip codes without data or with a 0 count should be displayed with a transparent color. The example above uses 5 shades of blue including the transparent color. You can either setup your color scale, or use `d3.scaleThreshold` with `d3.schemeBlues` that is available in the add-on D3's d3-scale-chromatic package. You can insert the following script after including D3 in your HTML code:

```
<script src="https://d3js.org/d3-scale-chromatic.v1.min.js"></script>
```

DATA SET – Two data sets are provided for this task:

nyc_zip.geojson: the boundaries for all zip codes in NYC. This is a GeoJSON file, where each feature only contains a zip code string (named "zipcode"), and the boundary for that zip code. For example:

```
[...
[-73.881919, 40.906666]]]],
    type: "MultiPolygon"
  },
  properties: [object Object] {
    zipcode: "10471"
  },
  type: "Feature"
}, ...]
```

nyc_restaurants_by_cuisine.json: this contains a list of restaurant distributions by zip code for each type of the cuisine. Each entry, which corresponds to a cuisine, consists of a "cuisine" name, a "total" field giving the total number of restaurants of that cuisine in NYC, and an associated array named "perZip" that has the count for each zip code with one or more restaurants of that cuisine. For example:

```
[...
[object Object] {
  cuisine: "Latin",
  perZip: [object Object] {
    10001: 67,
    10002: 111,
    10003: 52,
    10004: 59,
    ...
  },
  total: 17400
}
...]
```

The data sets are both available on NYU Classes, and from my GitHub repo:
https://raw.githubusercontent.com/hvo/datasets/master/nyc_zip.geojson
https://raw.githubusercontent.com/hvo/datasets/master/nyc_restaurants_by_cuisine.json

EXTRA CREDIT: +5 pts if you can add animation to your spatial visualization, i.e. when the user moves from cuisine to another, the map can smoothly transition its color distribution in half a second, for example.