

Lab 14.1 - Restaurant & Health Codes

Author Goes Here

Semester Goes Here

```
library(tidyverse)
```

The goal of this exercise is to visualize health code violations in New York City restaurants.

```
library(mdsr)
Violations
```

```
## # A tibble: 480,621 x 16
##   camis dba boro building street zipcode phone inspection_date action
##   <int> <chr> <chr>   <int> <chr>   <int> <dbl> <dtm>      <chr>
## 1  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2015-02-09 00:00:00 Viola~
## 2  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2014-03-03 00:00:00 Viola~
## 3  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-10-10 00:00:00 No vi~
## 4  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-09-11 00:00:00 Viola~
## 5  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-09-11 00:00:00 Viola~
## 6  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-08-14 00:00:00 Viola~
## 7  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-08-14 00:00:00 Viola~
## 8  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-08-14 00:00:00 Viola~
## 9  3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-08-14 00:00:00 Viola~
## 10 3.01e7 MORR~ BRONX    1007 MORRI~   10462 7.19e9 2013-08-14 00:00:00 Viola~
## # i 480,611 more rows
## # i 7 more variables: violation_code <chr>, score <int>, grade <chr>,
## #   grade_date <dtm>, record_date <dtm>, inspection_type <chr>,
## #   cuisine_code <dbl>
```

Following the instructions given in Section 18.6, Problem 1, modulo the following comments.

- 1.a. Geocoding
 - The Violations dataset has nearly a half-million records, with over 25K unique addresses. Geocoding that many addresses would likely take some hours to complete, so we suggest that you focus you filter down to those addresses that had more than 75 violations (or even more than that while you're developing your code).
 - Not all addresses geocode to actual coordinates, so plan to filter out any records with NA values for either lat or lon (nb. putting `na.omit()` into your pipeline will filter out all rows with NA values for any field.)
 - You'll note that the address information in this tibble is split over multiple fields. You can combine them into a workable address string using `paste(building, " ", street, " ", boro, " ", NY, " ", zipcode)`.
- 1.b. Static Mapping

- There a number of ways to illustrate the “nature and extent” of restaurant violations. We suggest starting the raw count of violations at each address. Feel free to go beyond this as time allows.
- 1.c. Interactive Mapping
 - We leave this interactive map step as an **optional** extension. As in lab 13, If you do that part, you’ll need to render your document using HTML (see the lab 13 specification for details on how to do this).