36-315 Final Project Data Pre-analysis

Arthur Jakobsson, Alex Cheng, Liz Chu, Kevin Ren

November 18, 2022

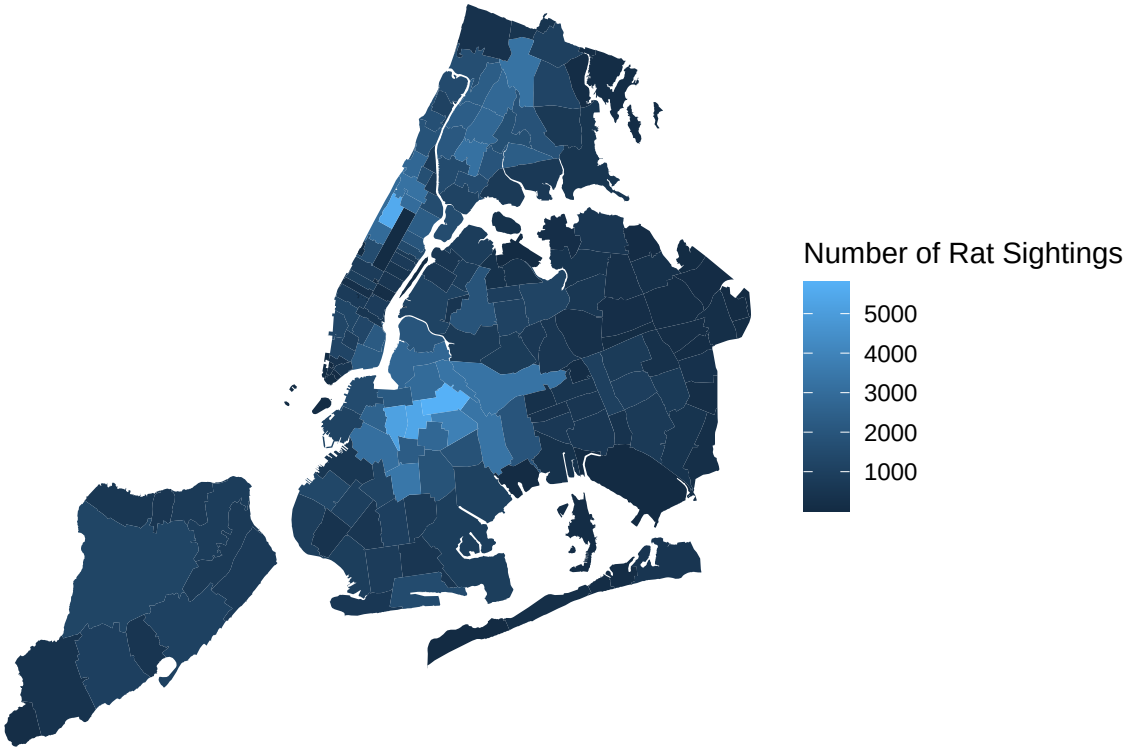
Contents

Rat Sighting Dataset - Data Pre-analysis

```
# library imports

library(tigris)
library(dplyr)
library(leaflet)
library(tidyverse)
library(sp)
library(ggmap)
library(maptools)
library(broom)
library(httr)
library(rgdal)
library(gridExtra)
library(stringr)
library(geosphere)
library(gpclib)
library(broom)
library(geojsonio)
```

```
ggplot() +
  geom_polygon(data = nyczips, aes(x = long, y = lat, group = group, fill = n_rats)) +
  theme_void() +
  # scale_fill_gradient2(low = "darkblue", mid = "purple", high = "pink", midpoint=3000) +
  coord_map() + labs(
    title = "Rat Sightings by Zip Code",
    fill = "Number of Rat Sightings"
  )
```

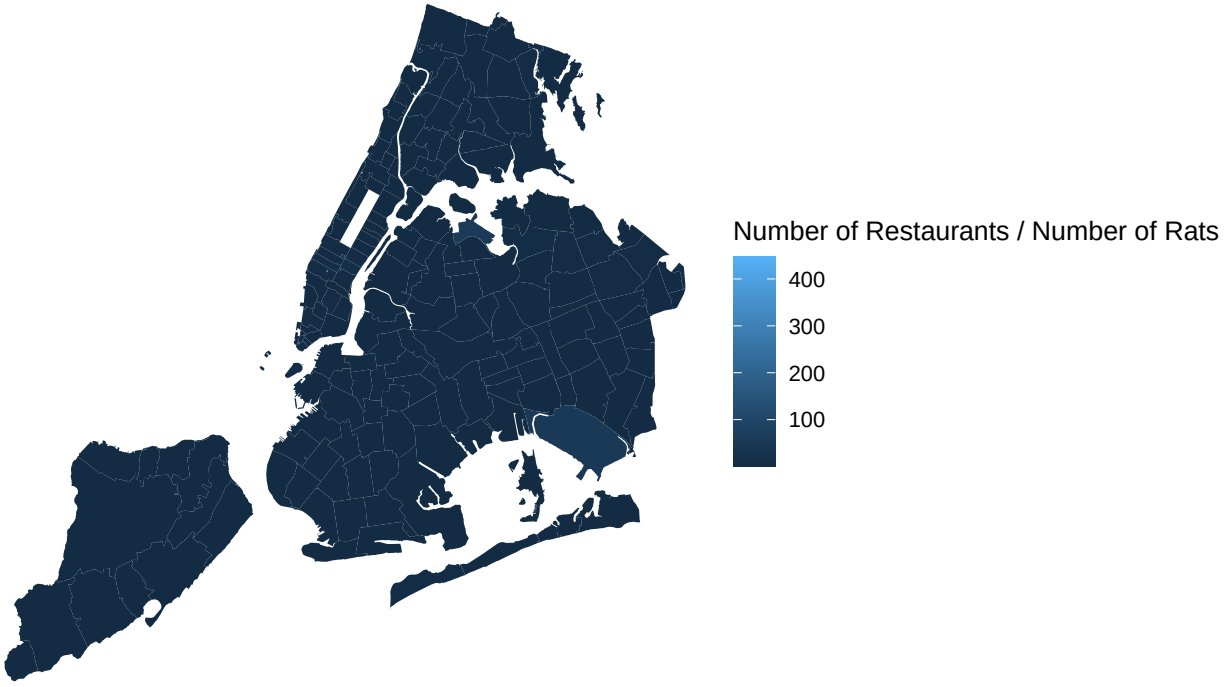
Rat Sightings by Zip Code




```
rest <- read.csv(file = "Raw/DOHMH_New_York_City_Restaurant_Inspection_Results.csv")
restsumzip = rest %>% group_by(ZIPCODE) %>% tally()

nyczips = nyczips %>% left_join(., restsumzip, by = c("id" = "ZIPCODE"))
colnames(nyczips)[9] = "n_rest"

nyczips = subset(nyczips, !is.na(n_rest))

nyczips$rest_to_rat = nyczips$n_rest / nyczips$n_rats
nyczips$rat_to_rest = nyczips$n_rats / nyczips$n_rest
```

```
par(mfrow=c(1,2))
ggplot() +
  geom_polygon(data = nyczips, aes(x = long, y = lat, group = group, fill = rest_to_rat)) +
  theme_void() +
  coord_map() + labs(
    title = "Restaurant to Rat Ratio by Zip Code",
    fill = "Number of Restaurants / Number of Rats"
  )
```

Restaurant to Rat Ratio by Zip Code




```
ggplot() +
  geom_polygon(data = nyczips, aes(x = long, y = lat, group = group, fill = rat_to_rest)) +
  theme_void() +
  coord_map() + labs(
    title = "Rat to Restaurant Ratio by Zip Code",
    fill = "Number of Rats / Number of Restaurants"
  )
```

Rat to Restaurant Ratio by Zip Code