

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

1

Rat Sighting Dataset - Data Pre-analysis

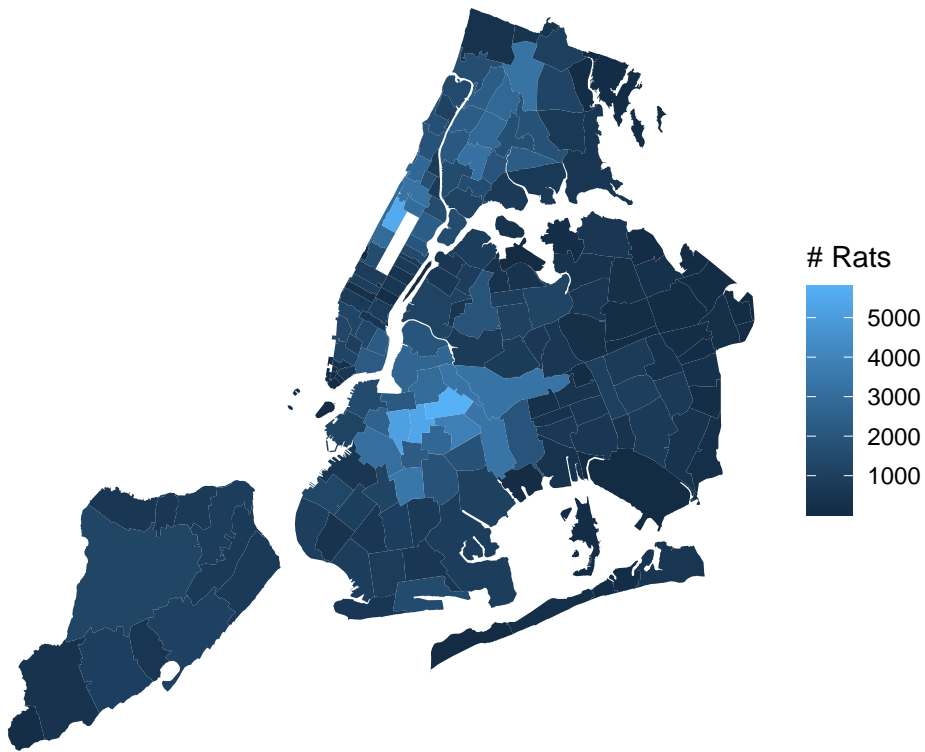
```
# library imports
```

```
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)
```

```
# SINGULAR VARIABLE ANALYSIS: # RATS, # RESTAURANTS, AVG RESTAURANT SCORE
```

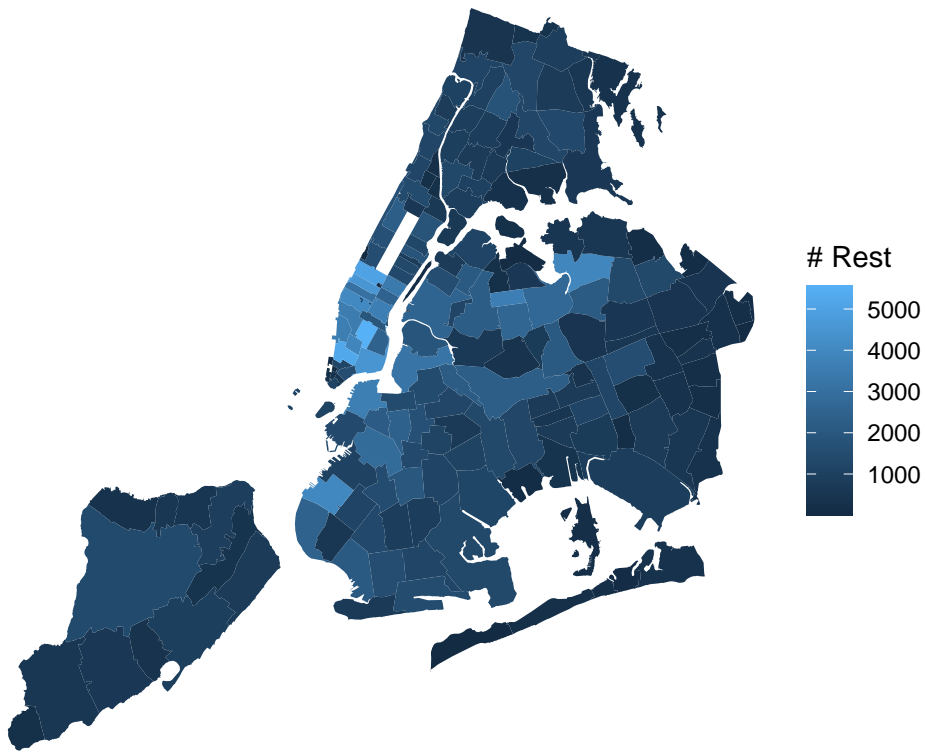
```
ggplot() +
  geom_polygon(data = nyczip, aes(x = long, y = lat, group = group, fill = n_rats)) +
  theme_void() +
  # scale_fill_gradient2(low = "darkblue", mid = "purple", high = "pink", midpoint=3000) +
  coord_map() + labs(
    title = "Rat Sightings by Zip Code",
    fill = "# Rats"
  )
```

Rat Sightings by Zip Code



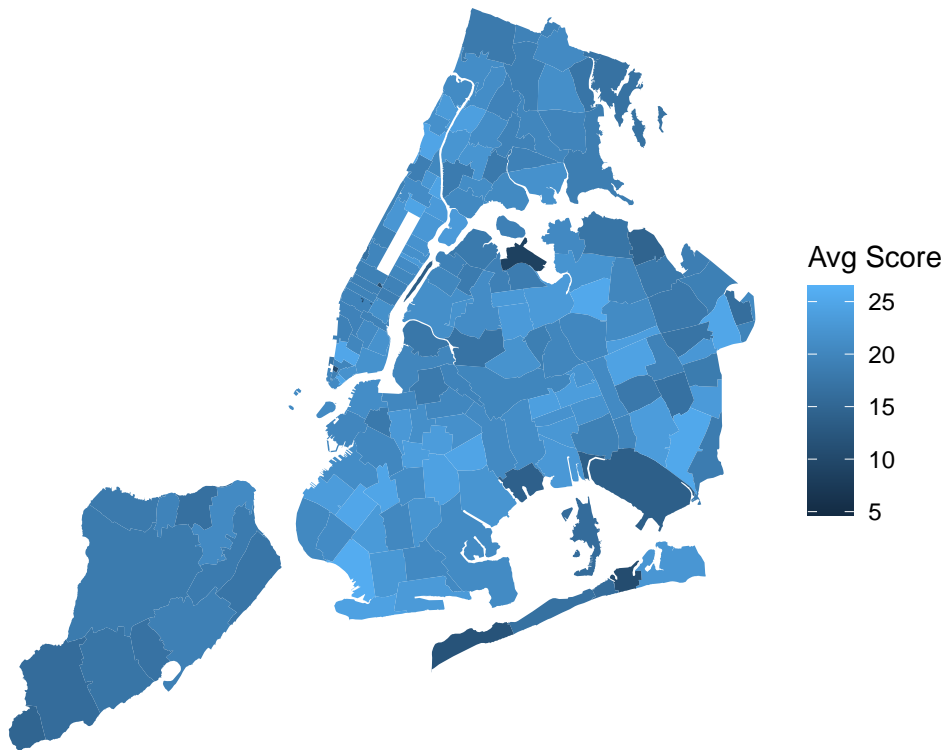
```
ggplot() +  
  geom_polygon(data = nyczip, aes(x = long, y = lat, group = group, fill = n_rest)) +  
  theme_void() +  
  coord_map() + labs(  
    title = "Number of Restaurants by Zip Code",  
    fill = "# Rest"  
  )
```

Number of Restaurants by Zip Code



```
ggplot() +  
  geom_polygon(data = nyczips, aes(x = long, y = lat, group = group, fill = score_avg)) +  
  theme_void() +  
  coord_map() + labs(  
    title = "Restaurant Scores by Zip Code",  
    fill = "Avg Score"  
  )
```

Restaurant Scores by Zip Code

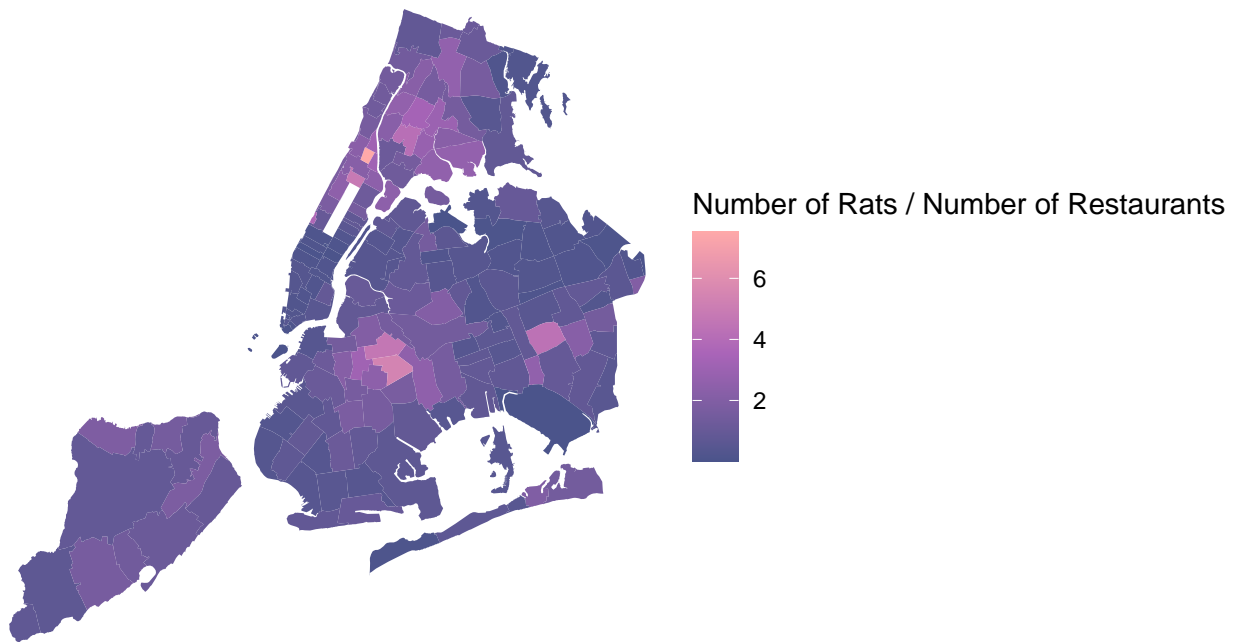


```
# THROWN OUT: RESTAURANT TO RAT RATIOS

# ggplot() +
#   geom_polygon(data = nyczip, aes(x = long, y = lat, group = group, fill = rest_to_rat)) +
#   theme_void() +
#   scale_fill_gradient2(low = "#395184",
#                         mid = "#A964B8",
#                         high = "#FFA9A9", midpoint = 200) +
#   coord_map() + labs(
#     title = "Restaurant to Rat Ratio by Zip Code",
#     fill = "Number of Restaurants / Number of Rats"
#   )

ggplot() +
  geom_polygon(data = nyczip, aes(x = long, y = lat, group = group, fill = rat_to_rest)) +
  theme_void() +
  scale_fill_gradient2(low = "#395184",
                      mid = "#A964B8",
                      high = "#FFA9A9", midpoint = 3.5) +
  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



```
# THROWN OUT: SCORE TO RAT RATIOS

# ggplot() +
#   geom_polygon(data = nyczip, aes(x = long, y = lat, group = group, fill = score_to_rat)) +
#   theme_void() +
#   scale_fill_gradient2(low = "#395184",
#                         mid = "#A964B8",
#                         high = "#FFA9A9", midpoint = 10) +
#   coord_map() + labs(
#     title = "Restaurant Score to Number of Rats by Zip Code",
#     fill = "Average Restaurant Score / Number of Rats"
#   )

ggplot() +
  geom_polygon(data = nyczip, aes(x = long, y = lat, group = group, fill = rat_to_score)) +
  theme_void() +
  scale_fill_gradient2(low = "#395184",
                      mid = "#A964B8",
                      high = "#FFA9A9", midpoint = 100) +
  coord_map() + labs(
    title = "Number of Rats to Restaurant Score by Zip Code",
    fill = "Number of Rats / Average Restaurant Score"
  )
```

Number of Rats to Restaurant Score by Zip Code

