STAT GR5702 EDAV Homework 4

Hantang Zhou (hz2562)

11/14/2018

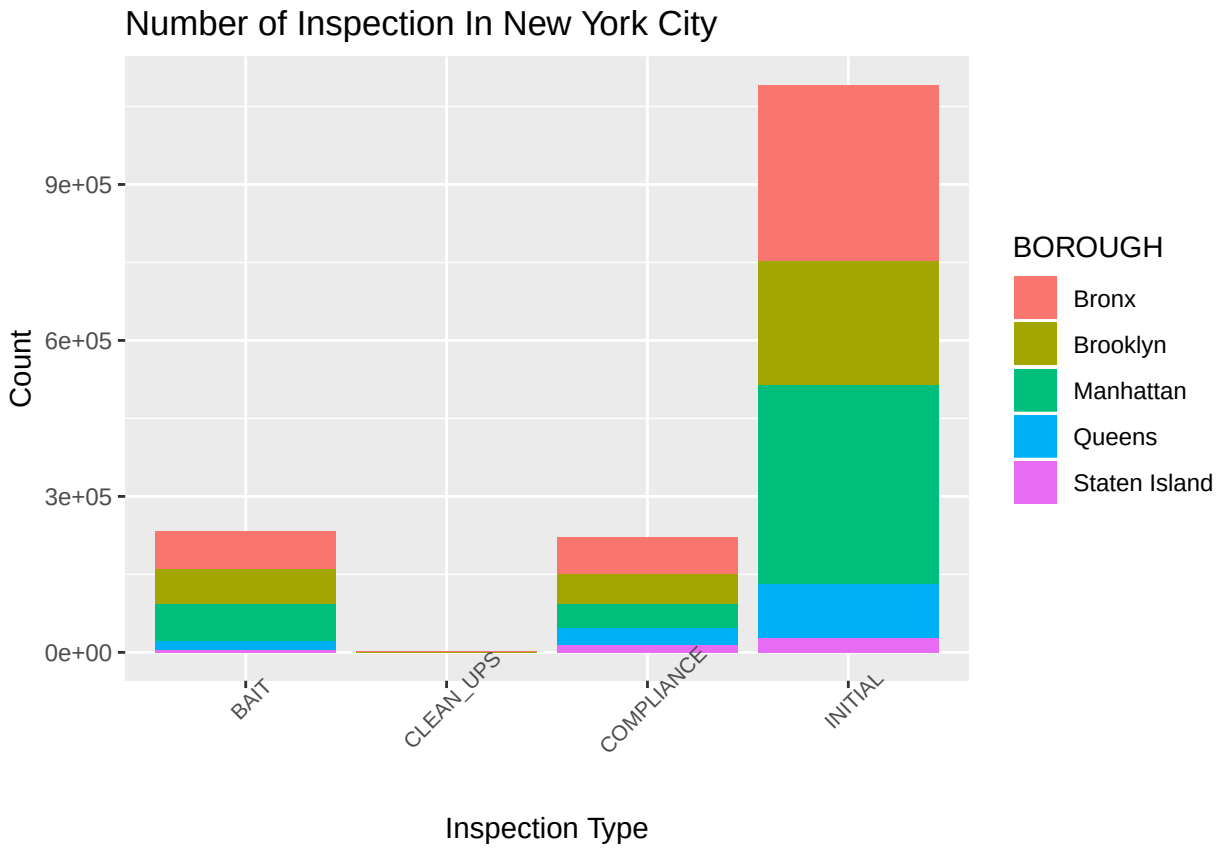
Data: Rodent Inspection

Sourse: NYCOpenData

Websites: https://data.cityofnewyork.us/Health/Rodent-Inspection/p937-wjvj

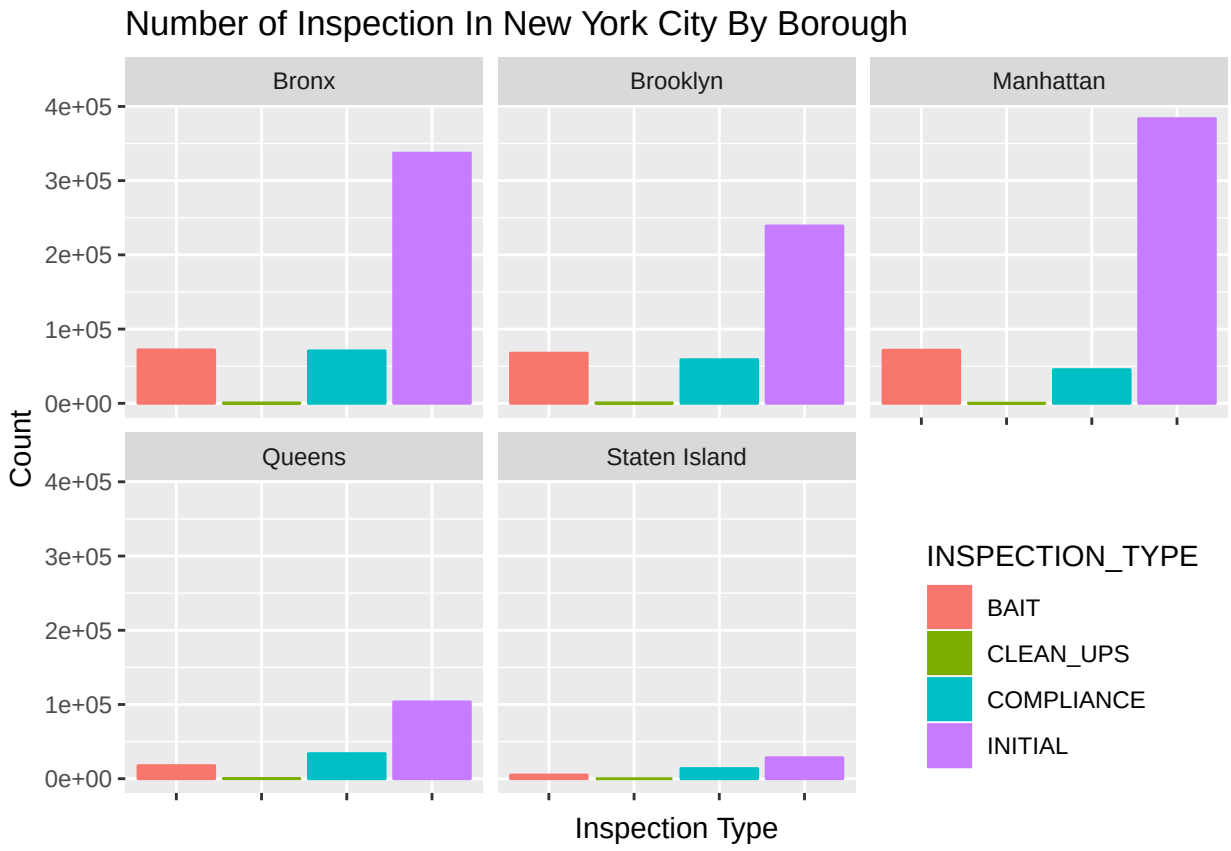
```
library(tidyverse)
setwd("/Users/hantangzhou/OneDrive/Columbia University in the City of New York/STAT GR5702/Final Projec
data <- read_csv('Rodent_Inspection.csv')

# bar chart of total inspection
bar_chart <- ggplot(data, aes(fill=BOROUGH, INSPECTION_TYPE)) + geom_bar() +
  theme(axis.text.x = element_text(size = 8, angle=45)) +
  ggtitle("Number of Inspection In New York City") + ylab("Count") + xlab("Inspection Type")
bar_chart
```




```
# facet bar chart
facet_bar_chart <- ggplot(data, aes(INSPECTION_TYPE, fill = INSPECTION_TYPE,
color = INSPECTION_TYPE)) + geom_bar() + facet_wrap(~BOROUGH) +
theme(axis.text.x = element_text(size = 0, angle=60), legend.position = c(0.85, 0.2)) +
ggtitle("Number of Inspection In New York City By Borough") + ylab("Count") +
```

```
  xlab("Inspection Type")
facet_bar_chart
```



Explaination and clarification:

- Initial Inspection - Inspection conducted in response to a 311 complaint, or a proactive inspection conducted through our neighborhood indexing program.
- Compliance Inspection - If a property fails its initial inspection, the Health Department will conduct a follow up (Compliance) inspection.
- Baiting - Application of rodenticide, or monitoring visit by a Health Department Pest Control Professional.
- Clean Up - The removal of garbage and clutter from a property by the Health Department.

Question raised:

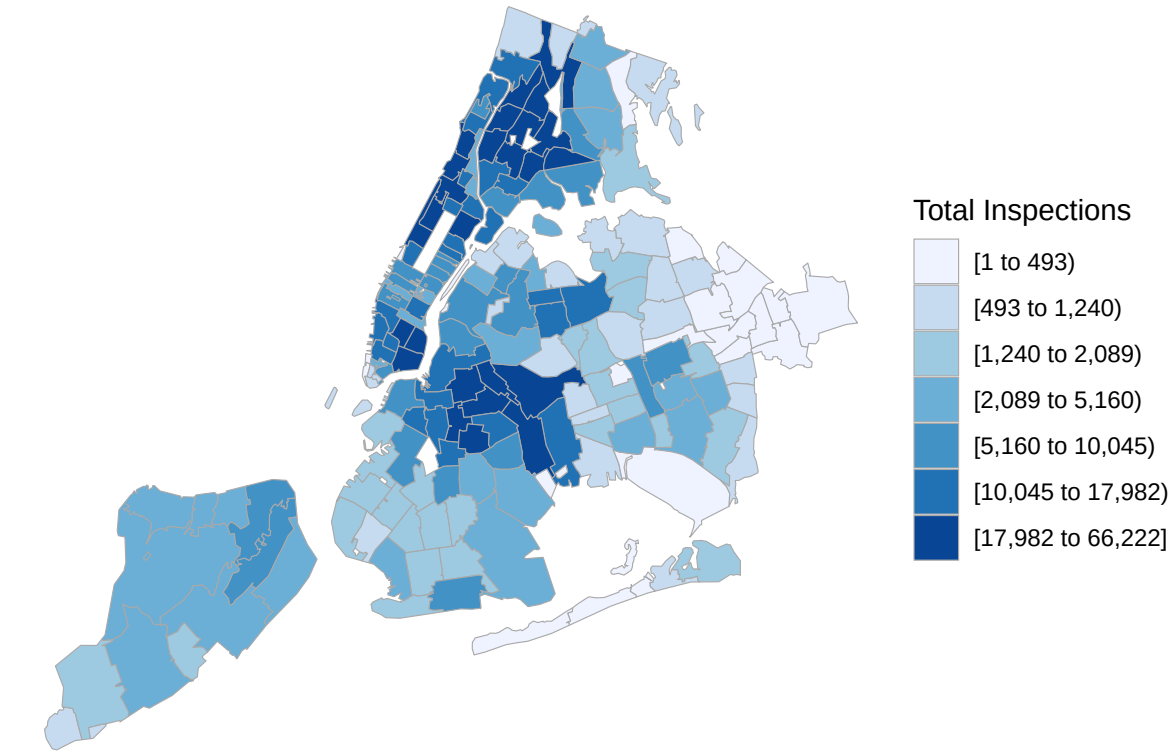
It seems that the numbers of different types of inspections are different. From the first graph, it seems that the compliance number of the bait nuber is about the same. According to the meaning of the inspection type, it means sense since after finding out there is a problem, then the bait is needed to be set.

However, from the second graph, it seems that for Manhattan, there are more "BAIT" than "COMPLIANCE", while in Queens it is the opposite. There should be a problem. I think we need to figure out how much taxlot/properties has been though the whole process ("INITIAL", "COMPLIANCE", "BAIT", the "CLEAN_UPS" is ignored as there are too few of clean-ups. I am gessing that the clean-ups are minor as no need for a field work).

```
# drawing the map by zip
library(choroplethr)
library(choroplethrMaps)
```

```
library(choroplethrZip)
data(zip.regions)

total_data <- data %>% group_by(ZIP_CODE) %>% summarize(value = n())
total_data <-  total_data %>% mutate(region = as.character(ZIP_CODE))
total_data <-  total_data %>% select(-ZIP_CODE)
total_data <- filter( total_data,  total_data$region %in% pull(zip.regions,1))
total_map <- zip_choropleth( total_data, zip_zoom =  total_data$region,
                             title      = "New York City Rodent Inspections By Zip",
                             legend     = "Total Inspections")
total_map
```

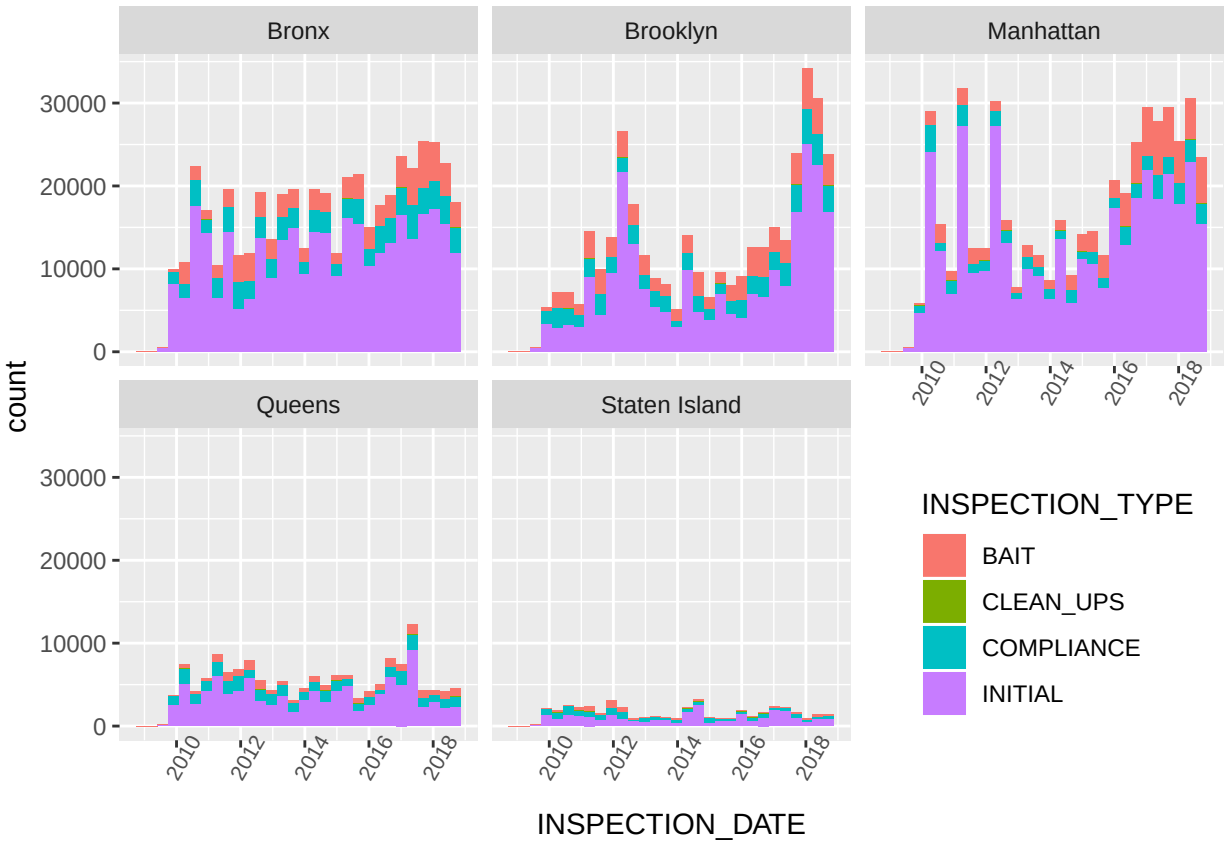



```
# represent the inspection by time
total_density <- ggplot(data=data, aes(INSPECTION_DATE, fill = INSPECTION_TYPE)) +
  geom_histogram()
total_density
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```


```
total_density_borough <- ggplot(data=data, aes(INSPECTION_DATE,
fill = INSPECTION_TYPE)) + geom_histogram() + facet_wrap(~BOROUGH) +
  theme(axis.text.x = element_text(size = 8, angle=60), legend.position = c(0.85, 0.2))
total_density_borough
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```



Question raised:

From the first chart, we can see there is a huge increase in inspections after 2016, and there is also a peak near between 2012 and 2013. It seems that different borough behave different under the same time period. It is worth the time to spend more time to examine the number of each type of inspections in differnt time for different borough.