

# NYC Sightings

Daniella Darkwah

2024-02-26

## Memo ‘Unveiling Brooklyn’s Rat Menace: A Tale of Urban Wildlife’

In the heart of New York City’s urban jungle, the persistent presence of rats has become a defining feature. Through a meticulous analysis of rat sightings data, a compelling narrative emerges, focusing particularly on the borough of Brooklyn. The objective is to uncover where and when rats are most prevalent and, more importantly, why Brooklyn takes the lead in this unsettling urban phenomenon.

Delving into the borough-specific data, Brooklyn consistently emerges as the epicenter of rat sightings. The narrative takes a thought-provoking turn as it aligns with the findings of an investigative article titled “Brooklyn Is the most Rat-Infested Borough, Locals Blame Construction and Gentrification.”(Hallowell, 2019). This in-depth exploration draws parallels between the ongoing construction projects and the gentrification wave sweeping across Brooklyn’s neighborhoods.

The article reveals a stark reality. The very development aimed at enhancing urban living inadvertently disrupts the delicate ecosystem, creating favorable conditions for rat infestations. As new residents flood into gentrifying areas, the construction boom displaces these resilient creatures, pushing them into the open, leading to an upsurge in rat sightings. This dual impact of construction and gentrification acts as a catalyst, making Brooklyn the borough most plagued by rats.

**A Visual Symphony of Contrast, Repetition, Alignment, and Proximity** In crafting the visual narrative, the principles of CRAP guide the presentation—Contrast, Repetition, Alignment, and Proximity. Contrast, vividly portrayed through distinctive colors, accentuates Brooklyn’s prominence amidst the five boroughs. Repetition ensures a consistent visual language, aiding comprehension and reinforcing key elements. Precise alignment and thoughtful proximity enhance the overall aesthetics and user experience, seamlessly guiding the audience through the data-rich narrative.

Alberto Cairo’s five qualities of great visualizations serve as the blueprint for this data-driven storytelling. Accuracy is paramount, as the visualizations faithfully represent the real data, fostering trust and reliability. Practicality governs the narrative flow, ensuring that the information is presented in a clear and accessible manner. Aesthetics, entwined with the principles of CRAP, elevate the visual experience, making it engaging and impactful.

Discernment is facilitated through strategic grouping and emphasis on key insights, allowing the audience to extract meaningful patterns. Illumination, the final quality, extends beyond mere visualization, shedding light on the complex interplay between urban development and the thriving rat population in Brooklyn.

**Unraveling the Complex Dance of Progress and Peril:** As the narrative unfolds, the exploration of rat sightings becomes a profound lens through which to understand the intricate dance between urban development and the flourishing rat population in Brooklyn. This tale not only informs but prompts a deeper reflection on the delicate balance between progress and the coexistence of city dwellers with their less-welcome neighbors. Through a rich interplay of data and storytelling, the audience is invited to contemplate the nuanced dynamics of urban wildlife, transcending the surface to unveil the profound implications of progress and peril in the heart of New York City.

## References

Hallowell, G. (2019, August 16). Brooklyn Is the most Rat-Infested Borough, Locals Blame Construction and Gentrification. *Bushwick Daily*. <https://bushwickdaily.com/news/6261-brooklyn-is-the-most-rat-infested-borough-locals-blame-construction-and-gentrification/>

## Loading Dataset

```
# Load necessary libraries
library(tidyverse)
library(readr)
library(dplyr)
library(lubridate)
```

```
# Read the dataset with custom NA values
raw_sightings <- read_csv("data/Rat_Sightings.csv", na = c("", "NA", "N/A"))
```

```
## Warning: One or more parsing issues, call 'problems()' on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)
```

```
## Rows: 101914 Columns: 52
## -- Column specification -----
## Delimiter: ","
## chr (33): Created Date, Closed Date, Agency, Agency Name, Complaint Type, De...
## dbl (6): Unique Key, Incident Zip, X Coordinate (State Plane), Y Coordinate...
## lgl (13): Landmark, Facility Type, School or Citywide Complaint, Vehicle Typ...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
# Check the first few rows to confirm successful loading and to view the structure
head(raw_sightings)
```

```
## # A tibble: 6 x 52
##   'Unique Key' 'Created Date'      'Closed Date'      Agency 'Agency Name'
##         <dbl> <chr>              <chr>          <chr>  <chr>
## 1  31464015 09/04/2015 12:00:00 AM 09/18/2015 12:00:00 ~ DOHMH Department o~
## 2  31464024 09/04/2015 12:00:00 AM 10/28/2015 12:00:00 ~ DOHMH Department o~
## 3  31464025 09/04/2015 12:00:00 AM <NA>          DOHMH Department o~
## 4  31464026 09/04/2015 12:00:00 AM 09/14/2015 12:00:00 ~ DOHMH Department o~
## 5  31464027 09/04/2015 12:00:00 AM 09/22/2015 12:00:00 ~ DOHMH Department o~
## 6  31464188 09/04/2015 12:00:00 AM 09/22/2015 12:00:00 ~ DOHMH Department o~
## # i 47 more variables: 'Complaint Type' <chr>, 'Descriptor' <chr>,
## #   'Location Type' <chr>, 'Incident Zip' <dbl>, 'Incident Address' <chr>,
## #   'Street Name' <chr>, 'Cross Street 1' <chr>, 'Cross Street 2' <chr>,
## #   'Intersection Street 1' <chr>, 'Intersection Street 2' <chr>,
## #   'Address Type' <chr>, 'City' <chr>, 'Landmark' <lgl>, 'Facility Type' <lgl>,
## #   'Status' <chr>, 'Due Date' <chr>, 'Resolution Action Updated Date' <chr>,
## #   'Community Board' <chr>, 'Borough' <chr>, ...
```

## Data Cleaning

```
raw_sightings <- raw_sightings %>%
  mutate(
    Created_Date = mdy_hms(`Created Date`),
    year = year(Created_Date),
    sighting_month = month(Created_Date),
    sighting_day = day(Created_Date),
    sighting_weekday = wday(Created_Date, label = TRUE)
  )
```

```
Sightings_summary <- raw_sightings %>%
  group_by(`Borough`, year) %>%
  summarize(number_of_sightings = n(), .groups = 'drop')
head(Sightings_summary)
```

```
## # A tibble: 6 x 3
##   Borough  year number_of_sightings
##   <chr>   <dbl>             <int>
## 1 BRONX   2010             2067
## 2 BRONX   2011             2174
## 3 BRONX   2012             2125
## 4 BRONX   2013             2120
## 5 BRONX   2014             2743
## 6 BRONX   2015             3189
```

```
# Rename key variables
raw_sightings <- raw_sightings %>%
  rename(
    unique_key = `Unique Key`,
    location_type = `Location Type`,
    incident_zip = `Incident Zip`,
    borough = `Borough`,
    latitude = `Latitude`,
    longitude = `Longitude`
  )
```

```
# Sort the boroughs by the total number of sightings and create a color palette
borough_order <- Sightings_summary %>%
  group_by(Borough) %>%
  summarize(Total_Sightings = sum(number_of_sightings)) %>%
  arrange(desc(Total_Sightings)) %>%
  pull(Borough)
```

```
my_pretty_theme <- theme_minimal(base_family = "Helvetica", base_size = 12) +
  theme(
    panel.grid.minor = element_blank(),
    plot.title = element_text(face = "bold", size = 16, hjust = 0.5), # Adjust the title size
    plot.subtitle = element_text(face = "plain", size = 12, color = "grey70", hjust = 0.5),
    plot.caption = element_text(face = "italic", size = 10, color = "grey70", hjust = 0),
    legend.title = element_text(face = "bold"),
```

```

strip.text = element_text(face = "bold", size = 12, hjust = 0),
axis.title = element_text(face = "bold", size = 12), # Same size as base to align with the text
axis.title.x = element_text(margin = margin(t = 10), vjust = 0.5), # Vertically center x-axis title
axis.title.y = element_text(margin = margin(r = 10), vjust = 0.5), # Vertically center y-axis title
strip.background = element_rect(fill = "grey90", color = NA),
panel.border = element_rect(color = "grey90", fill = NA)
)

# Arrange the Sightings_summary data frame in the same order
Sightings_summary$Borough <- factor(Sightings_summary$Borough, levels = borough_order)

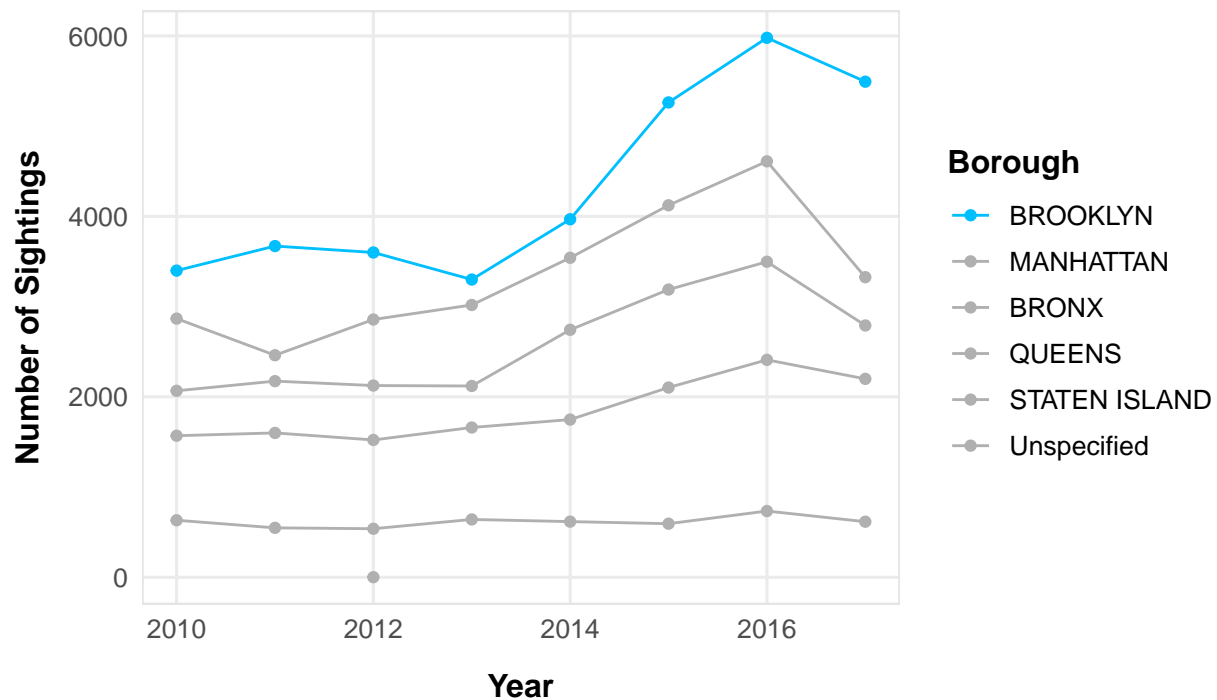
# Plot the data with the custom pretty theme
plot_1<-ggplot(Sightings_summary, aes(x = year, y = number_of_sightings, group = Borough, color = Borough)) +
  geom_line() + # Line plot
  geom_point() + # Add points to the line plot
  scale_color_manual(values = c("#00BFFF", "#B0B0B0", "#B0B0B0", "#B0B0B0", "#B0B0B0", "#B0B0B0")) +
  labs(
    title = "Highest Rat Sightings in New York City by Borough",
    subtitle = "Brooklyn leads in rat sighting over the years 2012 to 2017",
    x = "Year",
    y = "Number of Sightings",
    color = "Borough",
    caption = "Data source: NYC Health"
  ) +
  my_pretty_theme

plot_1

```

# Highest Rat Sightings in New York City by Borough

Brooklyn leads in rat sighting over the years 2012 to 2017



Data source: NYC Health

```
# Save the plot to a file
```

```
ggsave(plot_1, filename = "highest_rat_sightings.pdf", width = 10, height = 6)
```

```
# Arrange the Sightings_summary data frame in the same order
```

```
Sightings_summary$Borough <- factor(Sightings_summary$Borough, levels = borough_order)
```

```
# Plot the data with the custom pretty theme
```

```
plot_2 <- ggplot(Sightings_summary, aes(x = year, y = number_of_sightings, group = Borough, color = Borough))
```

```
  geom_line() + # Line plot
```

```
  geom_point() + # Add points to the line plot
```

```
  scale_color_manual(values = c("#00BFFF", "#B0B0B0", "#B0B0B0", "#B0B0B0", "#B0B0B0", "#B0B0B0")) +
```

```
  labs(
```

```
    title = "Highest Rat Sightings in New York City by Borough",
```

```
    subtitle = "Brooklyn leads in rat sighting over the years 2012 to 2017",
```

```
    x = "Year",
```

```
    y = "Number of Sightings",
```

```
    color = "Borough",
```

```
    caption = "Data source: NYC Health"
```

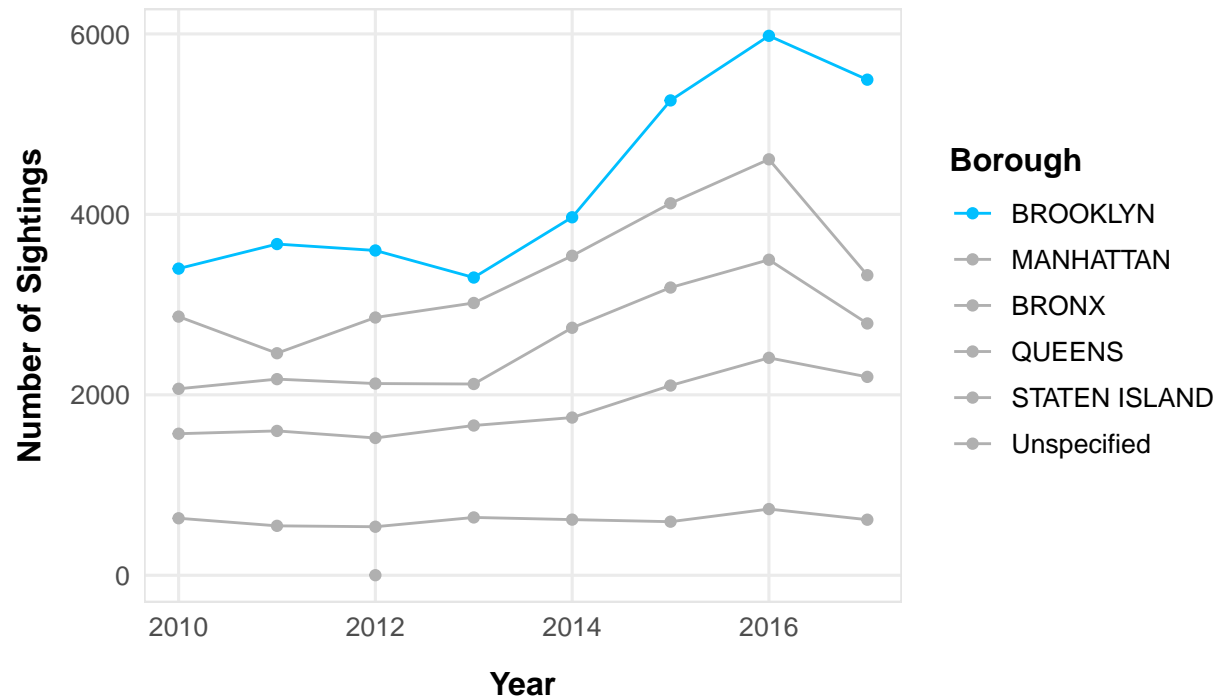
```
  ) +
```

```
  my_pretty_theme
```

```
plot_1
```

## Highest Rat Sightings in New York City by Borough

Brooklyn leads in rat sighting over the years 2012 to 2017



Data source: NYC Health

```
# Save the plot to a file
ggsave(plot_2, filename = "highest_rat_sightings.png", width = 10, height = 6)
```