New York City Film Permits Analysis

Mariya Tabachnikova

2022-12-18

Contents

Introduction

New York City has been the background for countless films and shows. The landmarks can be easily identified by most people and have become iconic symbols. Every year, thousands of permits are filed for permission to film at various locations across NYC all throughout the year. This report will explore any commonalities between filming frequency and time of year, borough, and more, in hopes of finding any trends and/or patterns.

R packages

```
library(tidyverse)
library(lubridate)
library(dplyr)
library(pander)
library(sf)
library(ggmap)
library(wesanderson)
```

Main Analysis

First, let's see how permit amount relates to time, in a day to day basis, on a monthly basis, and yearly.

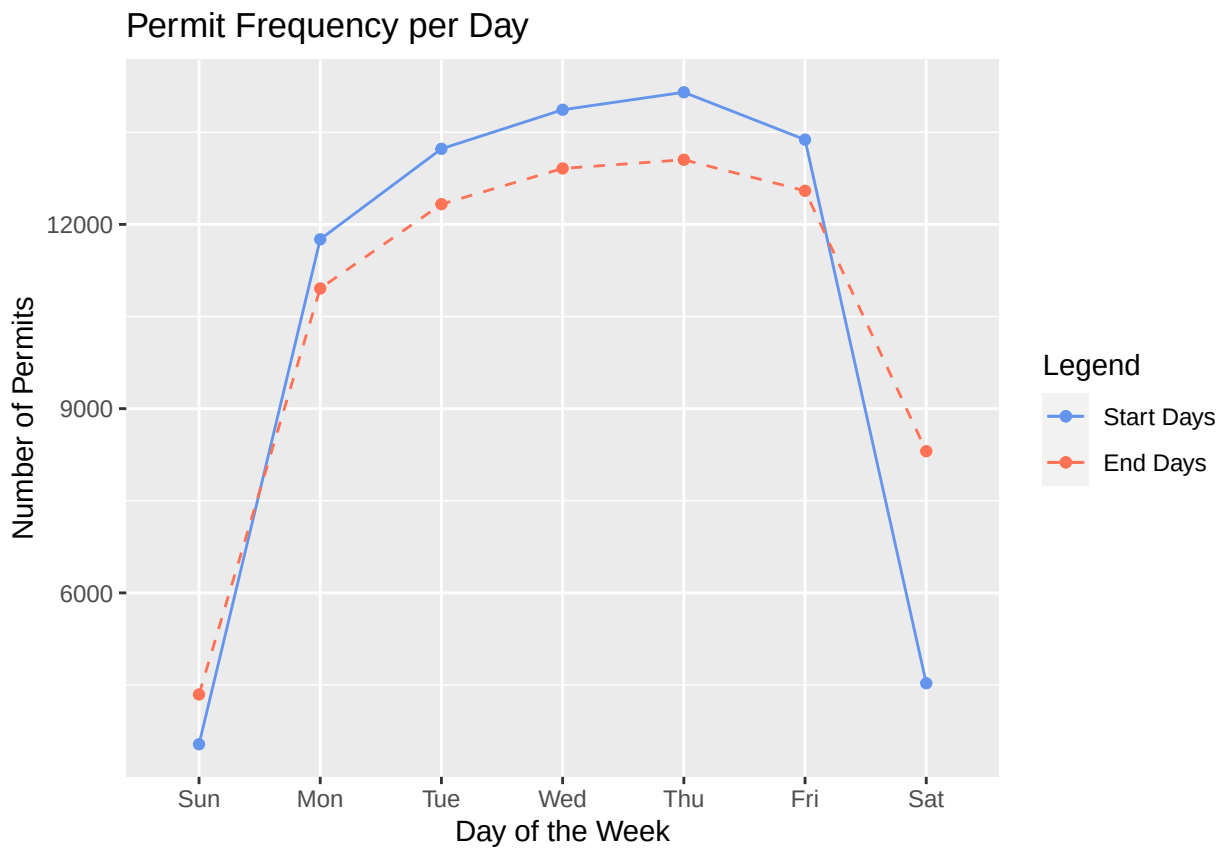
```
film_permits <- read_csv(file = "~/Documents/Film_Permits.csv")
```

```
film_permits <- film_permits %>%
  mutate( # parse the dates and
          # return them in standard date and time format
    StartDateTime = mdy_hms(StartDateTime),
    EndDateTime = mdy_hms(EndDateTime),
    EnteredOn = mdy_hms(EnteredOn)
    )

film_permits <- film_permits %>%
  mutate( # separate the dates and time into separate variables
          # so each variable can be called upon
    start_date = date(StartDateTime),
    end_date = date(EndDateTime),
    entered_date = date(EnteredOn),
    start_wday = wday(StartDateTime, label=TRUE),
    end_wday = wday(EndDateTime, label=TRUE),
    start_month = months(StartDateTime),
    start_year = isoyear(StartDateTime),
    start_time = hours(StartDateTime),
    end_time = hours(EndDateTime),
    entered_time = hours(EnteredOn)
    )
```

Permits Filed per Day

```
start_wday_tally <- film_permits %>%
                    group_by(start_wday) %>%
                    tally() # count permits started each weekday
end_wday_tally <- film_permits %>%
                    group_by(end_wday) %>%
                    tally() #count permits ended each weekday

ggplot() + geom_path(data = start_wday_tally,
                     aes(x = start_wday, y = n,
                         color = "cornflowerblue"),
                     group = 1) +
           geom_point(data = start_wday_tally,
                      aes(x = start_wday, y = n,   # plot permits started
                      color = "cornflowerblue")) + # per weekday
           geom_path(data = end_wday_tally,
                     aes(x = end_wday, y = n,
                         color="coral1"),
                     group = 1, linetype = "dashed") +
           geom_point(data = end_wday_tally,
                      aes(x = end_wday, y = n, # plot permits ended
```

```
                  color="coral1")) +       # per weekday on one graph
        ggtitle("Permit Frequency per Day ") +
        labs(x = "Day of the Week",
             y = "Number of Permits") +
        scale_color_identity(name = "Legend",
                             breaks = c("cornflowerblue","coral1"),
                             labels = c("Start Days", "End Days"),
                             guide = "legend")
```



It seems that filming during the middle of the week is most popular, with Thursday being the day when most productions start and end filming. The weekends are not popular filming days at all, but filming is more likely to end on Saturdays than it is to begin.

Permits Filed per Month

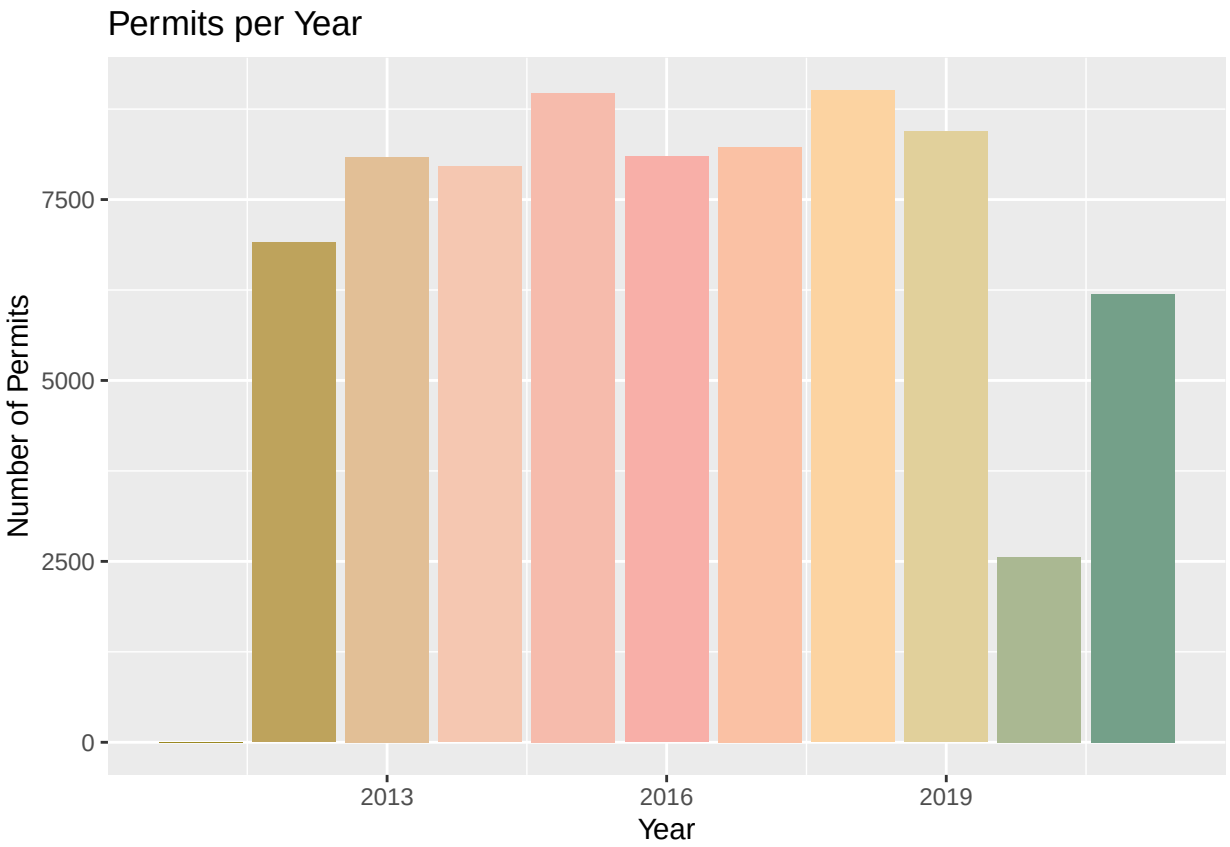
```
film_permits %>%
    mutate(Month = factor(start_month, # separate name of month from date variable
                          levels = month.name)) %>%
    arrange(Month) %>%
    group_by(Month) %>%
    tally() %>% # count permits per month
    rename("Number of Permits" = n) %>%
    pander()
```

Month	Number of Permits
January	4704
February	5276
March	6383
April	5591
May	5550
June	5485
July	6012
August	6848
September	6784
October	8588
November	7564
December	5663

Filming seems to take place more often in late summer to early autumn, with October having the most permits filed. As the temperatures drop and winter begins, filming drops significantly from November to January. It picks up again as winter ends in February and spring begins, before dropping again in early summer.

Permits Filed per Year

```
film_permits %>%
     group_by(start_year) %>%
     tally() %>%  # count number of permits filed each year
     ggplot(aes(x = start_year, y = n)) +
     labs(title = "Permits per Year",
          x = "Year",
          y = "Number of Permits") +
    geom_bar(stat = "identity",
             fill = wes_palette("Royal2", 11,  type = "continuous"))
```



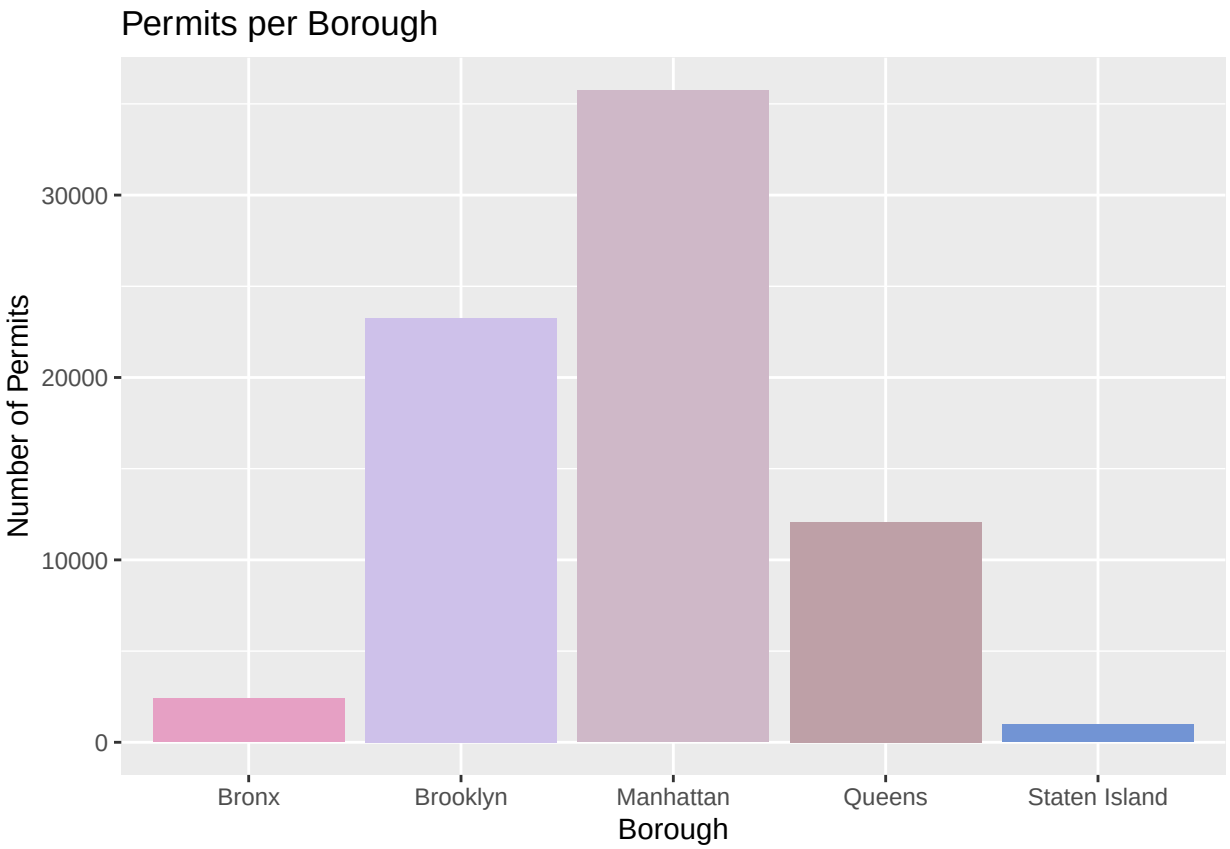
Through 2012 to 2019, the number of permits is pretty stable. However, it drops sharply in 2020. This can be explained by the COVID-19 pandemic, where everything but essential businesses were shut down. The number of permits has increased dramatically in 2021, but it was still not the same amount as before the pandemic. It clearly had an impact on the film industry, and it had not fully recovered in just one year's time.

Post Hoc Analysis

Now, let's look at how filming is related to location.

Permits Filed per Borough

```
film_permits %>%
  group_by(Borough)  %>%
  tally %>% # count how many permits filed per borough
  ggplot(aes(x = Borough, y = n ))  +
  labs(title = "Permits per Borough",
       y = "Number of Permits") +
  geom_bar(stat = "identity",
           fill = wes_palette("GrandBudapest2", 5,  type = "continuous"))
```



The bar graph shows that Manhattan is the most filmed borough out of all five, with Brooklyn in second, and Bronx and Staten Island falling far behind. This is not an unexpected result, as Manhattan is the most well-known borough, by far.

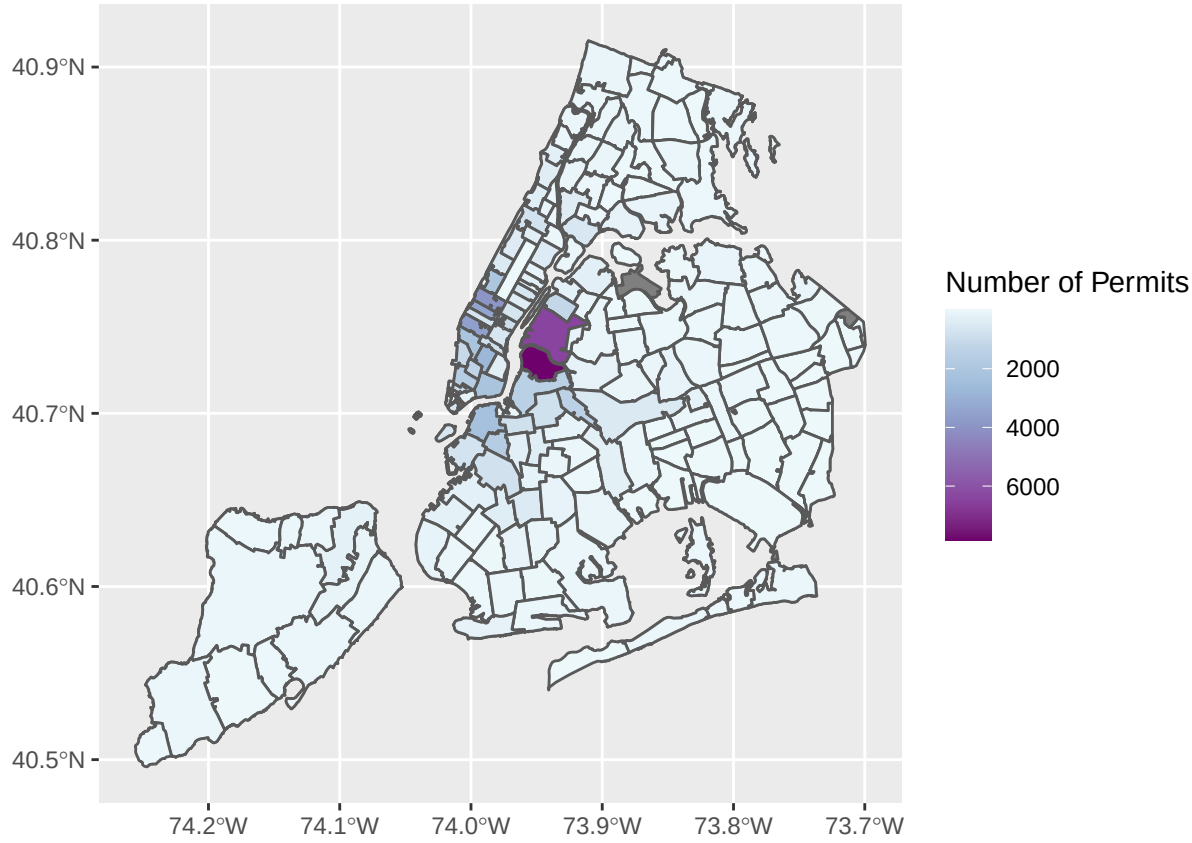
Permits Filed per ZIP Code

Now let's see if the most popular ZIP codes are all in Manhattan, too.

```
nyc_map <- st_read("~/Documents/archive/nyc-zip-code-tabulation-areas-polygons.geojson", quiet = TRUE)
# read in a map of NYC seperated by ZIP code lines

nyc_map <- st_transform(nyc_map, crs = 4326) # transform the map with spatial
                                             # referencing, so points can be
                                             # plotted accordingly
```

```
nyc_map <- nyc_map %>%
  rename(ZipCode = postalCode) %>% # rename variables to match for joining
  mutate(ZipCode = as.numeric(ZipCode)) # mutate to numeric because different
                                        # classes of variables cannot be joined

permit_count <- film_permits %>%
  separate_rows(ZipCode, sep = ",") %>% # separate multiple ZIP codes
                                        # from same column
  group_by(ZipCode) %>%
  summarise(N= n()) %>% # count how many times each ZIP code was filmed at
```

```
  mutate(ZipCode = as.numeric(ZipCode)) # mutate to numeric because different
                                        # classes of variables cannot be joined

permit_zipcodes <- left_join(nyc_map, permit_count, by = "ZipCode")
# join map by ZIP codes with permit count


ggplot(permit_zipcodes) + geom_sf() +
  aes(fill = N) + scale_fill_distiller(palette = "BuPu",
                                       trans = "reverse",
                                       name = "Number of Permits")
```



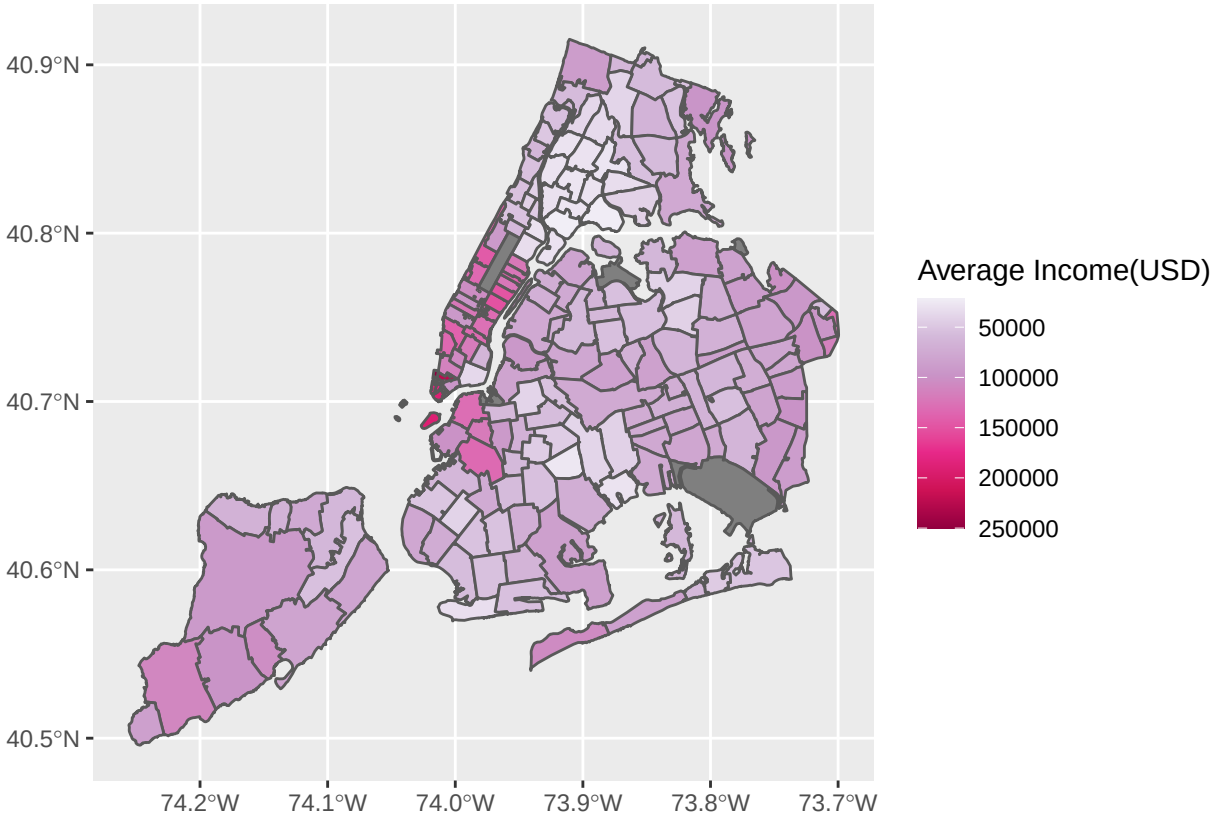
As the bar graph earlier showed, there is very little filming occurring in Staten Island and the Bronx. In Brooklyn, most filming takes places closer to Manhattan, like in Brooklyn Heights and Downtown Brooklyn. Most of the filming in Manhattan is in the lower half, under Central Park. Although Central Park is widely known, filming permits at parks cost more. This explains why there are not many filming permits being filed for Central Park. Most surprisingly, the highest frequency of filming by ZIP code takes place in Queens, specifically Long Island City, Ravenswood, and Astoria. When looking at a map, this makes more sense, as these locations would be able to have the famous NYC skyscrapers in the background.

Filed Permits vs Average Income

Let's see if filming locations are in any way related to average incomes of communities.

```
average_incomes <- read_csv(file = "~/Documents/average_incomes.csv")
# read in CSV containing average incomes of NYC by ZIP codes
```

```
income_zipcodes <- left_join(nyc_map, average_incomes, by = "ZipCode")
# both files already have the ZIP code variable to match, so they can be
# merged into one data set

ggplot(income_zipcodes) + geom_sf() + # plot incomes by ZIP code
aes(fill = AverageIncome) + scale_fill_distiller(palette = "PuRd",
                                                 trans = "reverse",
                                                 name = "Average Income(USD)")
```



According to this map, filming in Manhattan takes places most frequently in higher income neighborhoods, but not the richest ones. The most popular filming locations in Brooklyn are the ones that have a higher average income. However, the neighborhoods that have the most filming permits, which is in Queens, are not high earning income neighborhoods. Wealth of neighborhoods does not seem to be too closely related to popularity of filming, but it does occur rarely in low-income neighborhoods.

Conclusion

Filming does have very obvious patterns when it comes to the day of the week and the month. The middle of the week is more popular for filming, with a drop on weekends. Late summer to early autumn sees the most filming in NYC. On a yearly basis, filming has stayed pretty consistent and only faced a drastic change

because of the COVID-19 pandemic and the lock down it brought. The following year, 2021, did show that the film industry started to recover, but it had not yet reached its previous numbers.

When comparing boroughs, Manhattan had the most filming taking place, but the most sought after areas were in Queens, right across from Manhattan. When filming, the income level of a neighborhood is not the most important factor when choosing a location. In Manhattan, filming takes place in the lower part, which is not the most expensive neighborhood, but not the cheapest. For Brooklyn and Queens, it seems that income is less important than distance to Manhattan. Both boroughs' favorite spots were the ones straight across from Manhattan, where one could see skyscrapers in the distance.

Overall, filming seems to be centered around Manhattan, since locations are chosen based off proximity to this borough. Nonetheless, despite the after effects of COVID-19, NYC has been a consistently popular location for filming in the past decade and it looks like it will continue to be so, as the industry continues to grow once again.