Final

Rebecca Yu

1/28/2022

Introduction

According to the CDC "the risk for severe illness from COVID-19 increases with age," (CDC (2022)) this meaning that older generations infected with COVID-19 may require higher rates of intensive hospitalization and ventilation and higher rates of death. If nursing homes are a key setting where these older adults congregate and can spread disease, which New York State counties have the most and least fatalities within nursing homes when compared to the county's overall fatality count?

Using data from the US county COVID-19 data (Times (2022)) collected by The New York Times github repository and New York State nursing home data (Health (2022b)), I will find the New York State nursing homes with the highest and lowest fatality rates. I will then compare facility information (Health (2022a)) maintained by the New York State Department of Health in order to compare patterns in funding type via ownership, quality via inspection data, resident complaints, Medicare and Medicaid acceptance, and employee flu vaccination rate.

While none or all of these aspects of nursing home quality may or may not effect the rate of senior fatality as caused by COVID-19, I hope to determine whether the incidence in nursing homes with the highest fatalities can be reduced by altering aspects of the quality of care older adults receive at these facilities.

Packages Required

This project uses a majority of packages for standard data collection tidying, manipulation, and analysis. The one unique package used is the `urbnmapr` package. This package can create a map to visualize New York State counties via FIPs and is useful for data analysis.

```
suppressPackageStartupMessages({
    library(formatR)
    library(tidyverse)  # Makes tidy data
    library(stringr)  # Manipulates characters for tidy data
    library(knitr)  # Outputs data tables
    library(dplyr)  # Transforms and manipulates data
    library(readr)  # Reads .csv files and imports data
    library(ggplot2)  # Creates figures for visualizing data
    library(gridExtra)  # Visually compares figures
    library(devtools)  # Downloads packages from GitHub
    # devtools::install_github('UrbanInstitute/urbnmapr')
    library(urbnmapr)  # Creates U.S. county map
    library(gridExtra)  # Combines multiple plots
})
```

Data Preparation

The data for this project originated from the following sources:

- US county COVID-19 data
- New York State nursing home data
- New York State facility data

To begin my data preparation, I tidy three data sets by removing any unrelated columns and variables from `us_counties`, `Nursing_Home_Facility_Fatalities`, and `FACILITY_INFO`.

`us_counties` is an ongoing data set updated by the New York Times as a repository of data on COVID-19 cases and deaths in U.S counties. In this project, I take the data recorded before 01/11/2022. By filtering out any other `state` than New York. After this I remove the unofficial `City` from New York county and fill in its missing FIP. I also update the abbreviation `St. Lawrence` to `Saint Lawrence` in order to match the syntax in all data sets. It appears that the Bronx, Kings, Richmond, and Queens are missing from the `us.counties` data set and thus are set to missing. It is not clear if the New York Times had combined these data into the New York county or had just failed to include these counties in their data. Thus, I will assume they had failed to report on these counts and move forward assuming New York county counts for only New York county counts. I also separate this data into a separate data set for the day-to-day counts of COVID-19 fatalities. By doing this, I'm able to make a tidy data frame of New York State county COVID-19 fatalities by date in their corresponding county.

```
# Importing all data sets ----------------------

us_counties <- read_csv("/Users/rebecyu2/Desktop/Intermediate R/us-counties.csv")
Nursing_Home_Facility_Fatalities <- read_csv(paste0("/Users/",
    "rebecyu2", "/Desktop/", "Intermediate R/", "New_York_State",
    "_Statewide_", "COVID-19", "_Nursing_Home", "_and_Adult_",
    "Care_Facility", "_Fatalities.csv"))
FACILITY_INFO <- read_csv("/Users/rebecyu2/Desktop/Intermediate R/FACILITY_INFO.csv")

# Create tidy New York State county COVID-19 data
# ----------------------

nys_county_fatalities <- us_counties %>%
    filter(state == "New York") %>%
    select(date, county, fips, deaths) %>%
    rename(deaths_cum = deaths) %>%
    mutate(county = str_remove(county, ".City"))  # cumulative data

nys_county_fatalities_count <- nys_county_fatalities %>%
    group_by(county) %>%
    mutate(deaths = deaths_cum - lag(deaths_cum)) %>%
    mutate(deaths = replace(deaths, which(deaths < 0), 0))  #counts by date

summary(nys_county_fatalities)
```

```
##       date                county              fips             deaths_cum
##  Min.   :2020-03-01   Length:38616       Length:38616       Min.   :    0.0
##  1st Qu.:2020-08-30   Class :character   Class :character   1st Qu.:   12.0
##  Median :2021-02-13   Mode  :character   Mode  :character   Median :   64.0
##  Mean   :2021-02-12                                         Mean   :  722.4
```

```
##  3rd Qu.:2021-07-29                                    3rd Qu.:  151.0
##  Max.   :2022-01-11                                    Max.   :36089.0
```

```r
# I keep the cumulative `deaths` later on in this analysis
# since I will only use cumulative percentages when
# comparing counties and what's important to my analysis is
# the proportional differences between the counties #

nys_county_fatalities$fips[is.na(nys_county_fatalities$fips)] <- 36061
nys_county_fatalities$county[nys_county_fatalities$county ==
    "St. Lawrence"] <- "Saint Lawrence"
nys_county_fatalities_count$fips[is.na(nys_county_fatalities_count$fips)] <- 36061
nys_county_fatalities_count$county[nys_county_fatalities_count$county ==
    "St. Lawrence"] <- "Saint Lawrence"

# The New York Times data had been missing the Bronx,
# Queens, Kings, and Richmond counties that are present in
# the facility data#

head(nys_county_fatalities)
```

```
## # A tibble: 6 x 4
##   date        county      fips  deaths_cum
##   <date>      <chr>       <chr>      <dbl>
## 1 2020-03-01 New York    36061          0
## 2 2020-03-02 New York    36061          0
## 3 2020-03-03 New York    36061          0
## 4 2020-03-04 New York    36061          0
## 5 2020-03-04 Westchester 36119          0
## 6 2020-03-05 Nassau      36059          0
```

`Nursing_Home_Facility_Fatalities` is a data set maintained by the New York State Department of Health.
It reports the facility reported number of lab-confirmed COVID-19 fatalities in nursing homes for residents
that occurred at the facility and outside of the facility. I use the data reported before 01/12/2022. I clean
the data by separating the combined variable column of Department of Health Regional Office into `Office`
and `City`. I then mutate `Fac_County` and `Date` data and also remove the leading zeros in the `Fac_ID` to
match the format of `FACILITY_INFO` for later analysis. I also combine the counts of COVID-19 fatalities in
and out of the facility since I'm interested on the overall impact of nursing home facilities on U.S. adults
regardless of the location of the report. Thus, I'm able to make a tidy data frame of New York State county
COVID-19 fatalities in nursing homes by date in their corresponding county, city, and presiding office.

```r
# Create tidy New York State county nursing home COVID-19
# data ---------------------

nursinghome_fatalities <- Nursing_Home_Facility_Fatalities %>%
    separate(DOH_Reg_Office, c("Office", "City"), sep = ".-.") %>%
    mutate(Fac_County = str_to_title(Fac_County)) %>%
    mutate(Date = as.Date(Week_End_Report_Date, "%m/%d/%Y")) %>%
    rowwise() %>%
    mutate(Deaths = sum(c(Confirmed_In_Facility, Confirmed_Out_Of_Facility,
        Presumed_In_Facility))) %>%
    select(Office, City, Fac_County, Fac_ID, Fac_Name, Date,
        Deaths)
```

```
nursinghome_fatalities$Fac_ID <- str_remove(nursinghome_fatalities$Fac_ID,
    "^0+")
nursinghome_fatalities$Fac_ID <- as.double(nursinghome_fatalities$Fac_ID)
head(nursinghome_fatalities)
```

```
## # A tibble: 6 x 7
## # Rowwise:
##   Office          City    Fac_County Fac_ID Fac_Name         Date        Deaths
##   <chr>           <chr>   <chr>       <dbl> <chr>            <date>       <dbl>
## 1 Metropolitan ~ New Ro~ Westchest~     NA WILLOW TOWERS EHP 2020-04-21     11
## 2 Metropolitan ~ Long I~ Suffolk        NA VILLAGE WALK A CAR~ 2020-04-21    3
## 3 Metropolitan ~ New Yo~ Richmond       NA THE BRIELLE AT SEA~ 2020-04-21    4
## 4 Metropolitan ~ New Ro~ Westchest~     NA THE OSBORN        2020-04-21      3
## 5 Metropolitan ~ New Ro~ Westchest~     NA THE KENSINGTON    2020-04-21      5
## 6 Metropolitan ~ New Ro~ Westchest~     NA THE FOUNTAINS AT R~ 2020-04-21    8
```

FACILITY_INFO is a data set was created by the New York State Department of Health for their Nursing Home Profile public website as a monthly report of facility bed census, inspection results, complaint summaries, and state enforcement fine data. I use the data reported before 01/11/2022. I clean the data by separating the combined variable column of OWNERSHIP into OWNERSHIP an OWNERSHIP_TYPE. I'm more interested in the ownership of a nursing home whether it's proprietary or voluntary based in order to determine the general funding source. I also mutate the format of FACILITY NAME to match the format of Nursing_Home_Facility_Fatalities for later analysis. Thus, I'm able to make a tidy data frame of relevant nursing home facility information.

```
# Create tidy New York State county nursing home facility
# data ---------------------

facility_info <- FACILITY_INFO %>%
    separate(OWNERSHIP, c("OWNERSHIP", "OWNERSHIP_TYPE"), sep = "--") %>%
    select(FACILITY_NAME, FACILITY_ID, INSP_TOTAL, COMP_CITATIONS_PER_100,
        MEDICARE_CERTIFIED, MEDICAID_CERTIFIED, OWNERSHIP, EMPLOYEE_FLU_VACCINATION_RATE,
        AVERAGE_OCCUPIED_BEDS)
facility_info$FACILITY_NAME = str_to_upper(FACILITY_INFO$FACILITY_NAME)
facility_info$MEDICARE_CERTIFIED <- as.logical(facility_info$MEDICARE_CERTIFIED)
facility_info$MEDICAID_CERTIFIED <- as.logical(facility_info$MEDICAID_CERTIFIED)
head(facility_info)
```

```
## # A tibble: 6 x 9
##   FACILITY_NAME     FACILITY_ID INSP_TOTAL COMP_CITATIONS_P~ MEDICARE_CERTIF~
##   <chr>                  <dbl>      <dbl>            <dbl> <lgl>
## 1 ST. PETER'S NURSING~     17         23              1.3 TRUE
## 2 ST. MARGARET'S CENT~     19         31              2.5 TRUE
## 3 HUDSON PARK REHABIL~     21         24              5.8 TRUE
## 4 DAUGHTERS OF SARAH ~     22         18              3.9 TRUE
## 5 TERESIAN HOUSE NURS~     23         38              2   TRUE
## 6 DELMAR CENTER FOR R~     27         65             13.9 TRUE
## # ... with 4 more variables: MEDICAID_CERTIFIED <lgl>, OWNERSHIP <chr>,
## #   EMPLOYEE_FLU_VACCINATION_RATE <dbl>, AVERAGE_OCCUPIED_BEDS <dbl>
```

Exploratory Data Analysis

The primary goal of this analysis is to determine differences in nursing home quality by the highest and lowest COVID-19 fatalities in New York State nursing homes by county. Therefore, I will begin by calculating the counties with the highest and lowest nursing home fatalities and then assess the differences in quality these sets of nursing homes divulge.

New York State counties with the highest and lowest nursing home fatalities

```
# Calculations for the percent of nursing home fatalities
# over the overall county fatalities due to COVID-19
# --------------------

# * Initial calculations comparing day-to-day fatality
# counts --------------------- *
combined_fatalities <- nys_county_fatalities_count %>%
    select(county, fips, deaths, date) %>%
    right_join(nursinghome_fatalities, by = c(county = "Fac_County",
        date = "Date"))

# I combine the overall county fatalities to nursing home
# fatalities by date and county #

Delaware <- combined_fatalities %>%
    select(county, date, deaths, Deaths) %>%
    filter(county == "Delaware")
head(Delaware)   # recorded data post 2020
```

```
## # A tibble: 6 x 4
## # Groups:   county [1]
##   county   date        deaths Deaths
##   <chr>    <date>       <dbl>  <dbl>
## 1 Delaware 2020-12-15       0      1
## 2 Delaware 2020-12-29       0      2
## 3 Delaware 2021-01-05       0      3
## 4 Delaware 2021-01-12       0      3
## 5 Delaware 2021-01-19       0     10
## 6 Delaware 2021-01-26       0      4
```

```
combined_fatalities %>%
    select(county, date, deaths, Deaths) %>%
    filter(county == "Albany") %>%
    rename(NYT_deaths = deaths) %>%
    rename(facility_deaths = Deaths) %>%
    pivot_longer(cols = c(NYT_deaths, facility_deaths), names_to = "Fatalities",
        values_to = "Count") %>%
    ggplot(aes(x = date)) + geom_line(mapping = aes(y = Count,
    group = Fatalities, color = Fatalities)) + labs(title = paste0("Supplemental",
    " Figure 1:, 2020 fatality", " data inconsistency"))  # recorded data in 2020
```

Supplemental Figure 1:, 2020 fatality data inconsistency
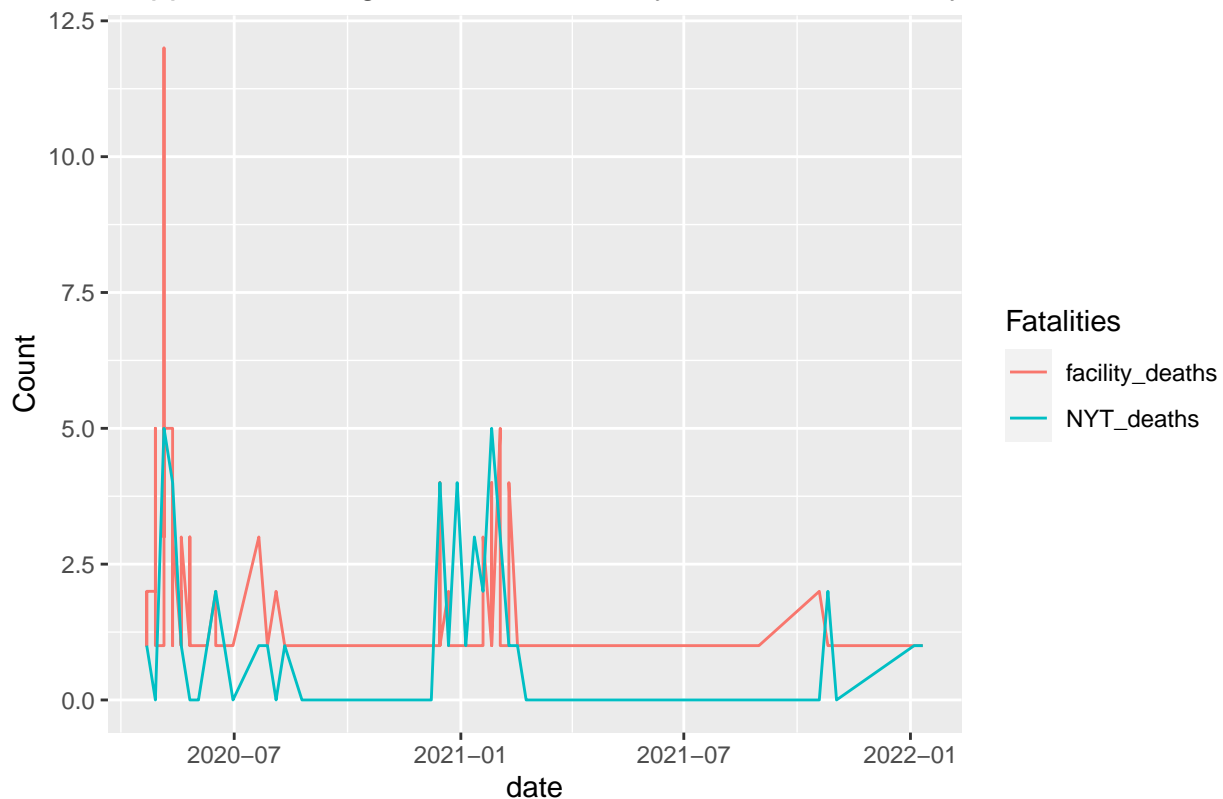

```
combined_fatalities <- nys_county_fatalities_count %>%
    select(county, fips, deaths, date) %>%
    right_join(nursinghome_fatalities, by = c(county = "Fac_County",
        date = "Date")) %>%
    separate(date, c("year", "date"), sep = 4, convert = TRUE) %>%
    select(county, fips, deaths, Deaths, year, date, Fac_Name,
        Fac_ID) %>%
    filter(year != 2020) %>%
    unite(date, year, date, sep = "") %>%
    drop_na()

percent_fatalities <- combined_fatalities %>%
    group_by(county, fips) %>%
    drop_na() %>%
    filter(sum(deaths) > sum(Deaths)) %>%
    summarize(Percent = round((sum(Deaths)/sum(deaths)) * 100,
        digits = 2)) %>%
    ungroup()  # Percent according to day by day count data

# * New calculations using cumulative data in order to view
# proportional trends ----------------------
combined_fatalities <- nys_county_fatalities %>%
    select(county, fips, deaths_cum, date) %>%
    right_join(nursinghome_fatalities, by = c(county = "Fac_County",
        date = "Date"))

percent_fatalities <- combined_fatalities %>%
```

```
    group_by(county, fips) %>%
    filter(sum(deaths_cum) > sum(Deaths)) %>%
    summarize(Percent = round((sum(Deaths)/sum(deaths_cum)) *
        100, digits = 2)) %>%
    ungroup() %>%
    drop_na()  # Percent according to cumulative data
```

We can see that there are dates post-2020 that the New York Times have counted no deaths due to COVID-19, however the nursing homes have counted deaths (for example: if we look at the data in Delaware county). This discrepancy may relate to a difference in the definition of what counts a death due to COVID-19.

If we look at the dates during the year 2020 (for example: in Albany), we can see that there's a way larger number in the facility deaths than counted by the New York Times, thus making the proportion of county deaths to nursing home deaths impossible. We can see that post 2020, the peaks become similar. Therefore, I remove any data from before 2021 from my data set.

Now, only 10 counties have NYT total county deaths that are greater than nursing home total county deaths. In order to move forward with my data analysis, I decide to use the initial cumulative data from the New York Times, over comparing date to date, in order to get a simple proportion for all counties. This is because I'm interested in observing a trend in the data in order to get the bigger picture and best representation of New York State.

With this new data set with each county's percent of nursing home COVID-19 fatalities out of total county COVID-19 fatalities, I will visualize the difference between the counties with the highest and lowest percentage of nursing home fatalities due to COVID-19.

```
# Counties with the highest fatalities
# ---------------------

top_five <- percent_fatalities %>%
    arrange(desc(Percent)) %>%
    slice(1:5)  # Stored as a data frame
top_five_char <- top_five$county  # Stored as a character vector

# Counties with the lowest fatalities
# ---------------------

last_five <- percent_fatalities %>%
    arrange(desc(Percent)) %>%
    slice(53:57)  # Stored as a data frame
last_five_char <- last_five$county  # Stored as a character vector

# * Compilation of highest and lowest counties
# --------------------- *
highlight_counties <- rbind(top_five, last_five)  # Compilation of counties I'm focusing on
highlight_counties_char <- append(top_five_char, last_five_char)  # Compilation for factoring

# Figure 1: Highest and lowest nursing home fatalities due
# to COVID-19 ---------------------

ggplot(highlight_counties) + geom_bar(mapping = aes(x = factor(county,
    levels = highlight_counties_char), y = Percent, fill = Percent),
    stat = "identity") + labs(x = "County", title = paste0("Figure 1: Highest",
    " and lowest", " nursing home", " fatalities due to COVID-19")) +
    theme(axis.text.x = element_text(angle = 45, hjust = 1))
```
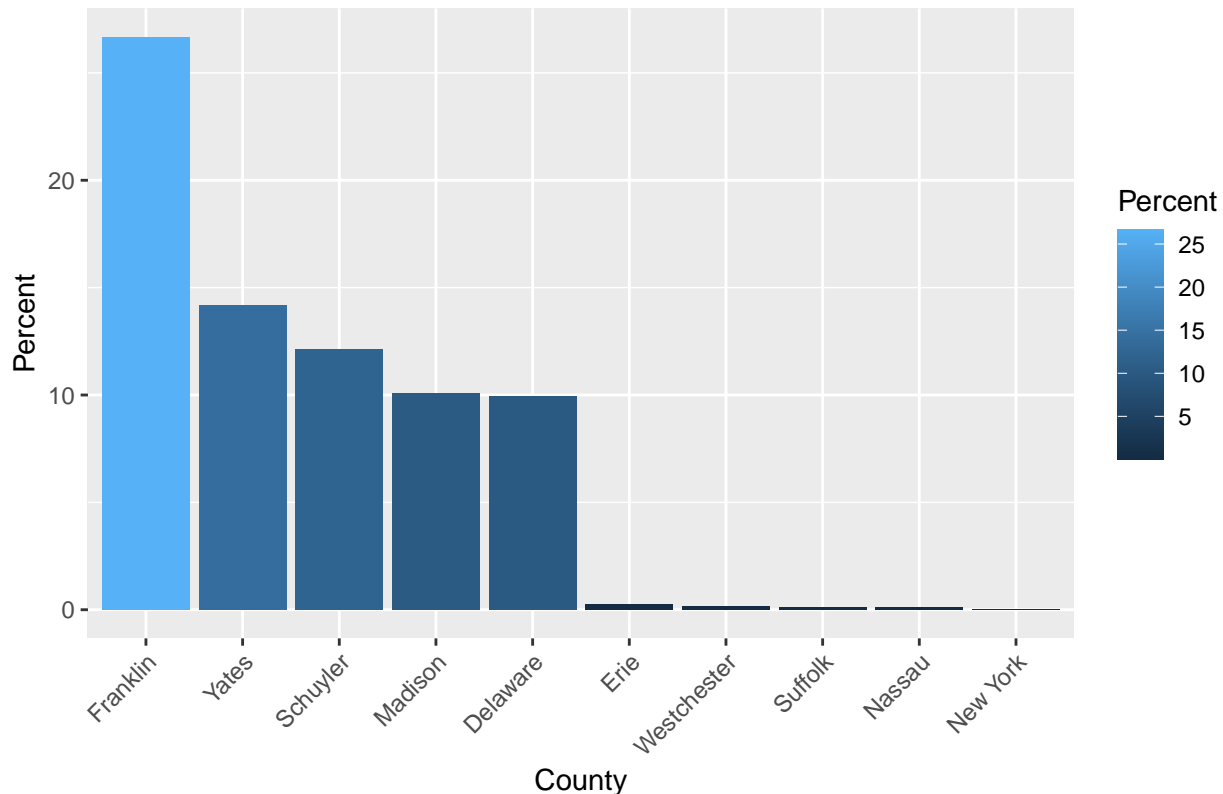
Figure 1: Highest and lowest nursing home fatalities due to COVID−19

It's evident in Figure 1 that the counties with the highest percentage of fatalities are vastly different from those with the lowest. Now that I have an idea of this disparity, I will make a heat map out of the New York State counties to visualize any possible geographic patterning.

```
# Figure 2: Heat map of New York state county nursing home
# fatalities ---------------------
countydata <- countydata %>%
    left_join(counties, by = "county_fips") %>%
    filter(state_name == "New York")
countydata <- countydata %>%
    left_join(percent_fatalities, by = c(county_fips = "fips"))
countydata %>%
    ggplot(mapping = aes(long, lat, group = group, fill = Percent)) +
    geom_polygon(color = "#ffffff", size = 0.25) + coord_map(projection = "albers",
    lat0 = 39, lat1 = 45) + scale_fill_gradientn(colors = c("#0032fc",
    "#f54275")) + theme(legend.title = element_text(), legend.key.width = unit(0.5,
    "in")) + labs(fill = "Percentage", title = paste0("Figure 2:",
    " Heat map", " of New York state", " county nursing home fatalities"))
```

Figure 2: Heat map of New York state county nursing home fatalities
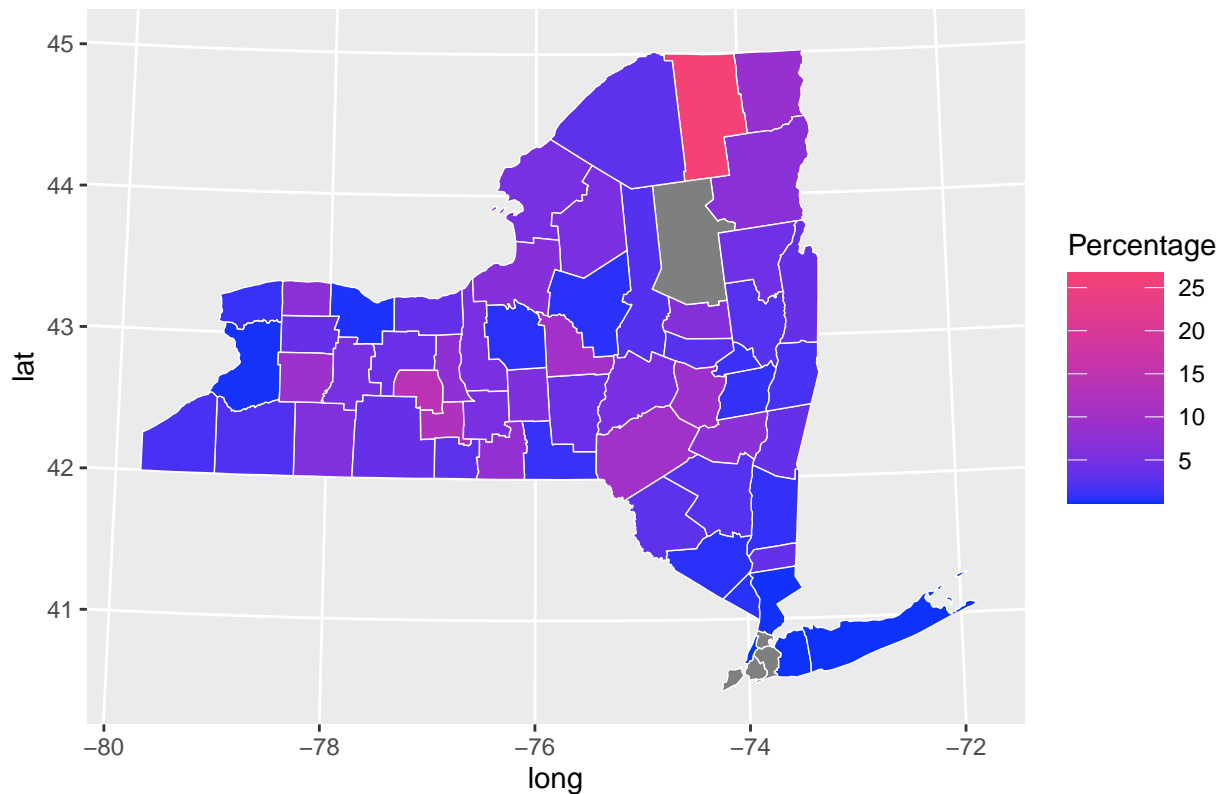
While a clear geographic distinction is not evident, we can see a slight trend in lower COVID-19 fatalities in counties that encompass larger cities as opposed to the larger, more central counties which lean into a purple-red. Overall, Figure 2 gives us a sense that the cultures of the geographic regions of New York State isn't the greatest factor in nursing home fatalities.

Analyzing differences in nursing home quality by the highest and lowest COVID-19 fatalities in New York State nursing homes

```
# Compilation of the specific nursing homes with the
# counties with the highest and lowest fatalities
# ----------------------

top_five_nursing <- top_five %>%
    left_join(nursinghome_fatalities, by = c(county = "Fac_County"))
last_five_nursing <- last_five %>%
    left_join(nursinghome_fatalities, by = c(county = "Fac_County"))

# Compilation of the highest and lowest nursing home
# fatality percentages and their facility information by
# county ----------------------

top_five_nursing <- top_five_nursing %>%
    inner_join(facility_info, by = c(Fac_ID = "FACILITY_ID",
        Fac_Name = "FACILITY_NAME")) %>%
    select(county, Percent, Fac_Name, INSP_TOTAL, COMP_CITATIONS_PER_100,
```

```
            MEDICARE_CERTIFIED, MEDICAID_CERTIFIED, OWNERSHIP, EMPLOYEE_FLU_VACCINATION_RATE,
            AVERAGE_OCCUPIED_BEDS)

last_five_nursing <- last_five_nursing %>%
    inner_join(facility_info, by = c(Fac_ID = "FACILITY_ID",
        Fac_Name = "FACILITY_NAME")) %>%
    select(county, Percent, Fac_Name, INSP_TOTAL, COMP_CITATIONS_PER_100,
        MEDICARE_CERTIFIED, MEDICAID_CERTIFIED, OWNERSHIP, EMPLOYEE_FLU_VACCINATION_RATE,
        AVERAGE_OCCUPIED_BEDS)

highlight_nursing <- rbind(top_five_nursing, last_five_nursing)
```

After creating these new data sets with the facility information of the nursing homes within the counties with the highest and lowest fatalities, I used bar graphs to show the differences in multiple facility aspects including quality via inspection data, resident complaints, Medicare and Medicaid acceptance, and employee flu vaccination rate. These data is under Supplemental Figures.

First, let's look at how many nursing homes are within each county.

```
# Figure 3: Number of nursing homes per county
# ----------------------

quantity_nh <- highlight_nursing %>%
    group_by(county) %>%
    summarise(Quantity = length(Fac_Name))

ggplot(quantity_nh) + geom_bar(mapping = aes(x = factor(county,
    levels = highlight_counties_char), y = Quantity, fill = Quantity),
    stat = "identity") + labs(x = "County", y = "Count", title = paste0("Figure 3:",
    " Number of ", "nursing homes", " per county")) + theme(axis.text.x = element_text(angle = 45,
    hjust = 1))
```

Figure 3: Number of nursing homes per county
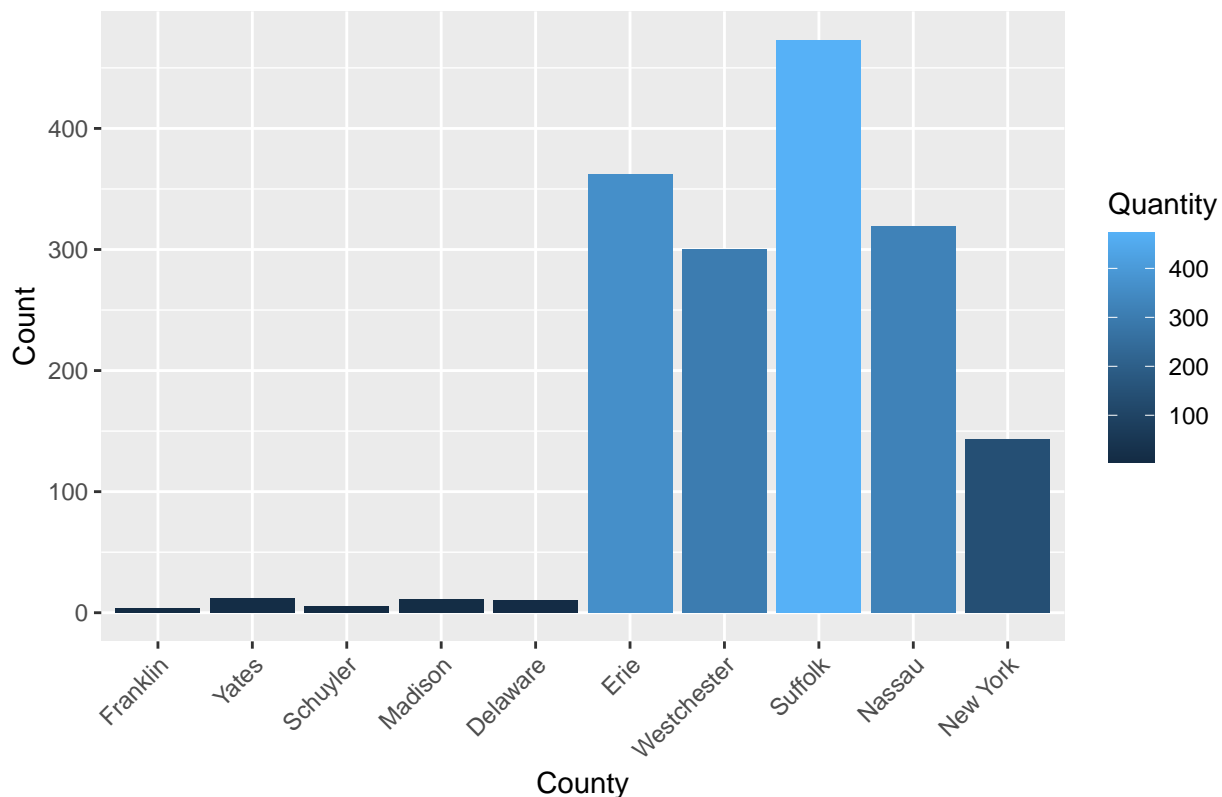
By looking at the number of nursing homes in the counties with the highest fatalities, we can see it's significantly less than that of the counties with the least number of fatalities. It's interesting to note as one would assume with more nursing homes, there should be more fatalities in nursing homes per overall COVID-19 deaths.

Next, another notably significant difference between nursing homes with the highest and lowest fatalities is see in funding type.

```
# Figure 4: Funding sources for the highest and lowest
# nursing home fatalities ----------------------

top_funding_plot <- ggplot(data = top_five_nursing) + geom_bar(mapping = aes(x = OWNERSHIP,
    fill = factor(county, levels = top_five_char))) + labs(y = "Count",
    x = "Ownership", fill = "County") + theme(axis.text.x = element_text(angle = 45,
    hjust = 1))

last_funding_plot <- ggplot(data = last_five_nursing) + geom_bar(mapping = aes(x = OWNERSHIP,
    fill = factor(county, levels = last_five_char))) + labs(y = "Count",
    x = "Ownership", fill = "County") + theme(axis.text.x = element_text(angle = 45,
    hjust = 1))

grid.arrange(top_funding_plot, last_funding_plot, nrow = 1, top = paste0("Figure 4:",
    " Funding sources", " for the highest", " and lowest nursing",
    " home fatalities"))
```

Figure 4: Funding sources for the highest and lowest nursing home fatalities
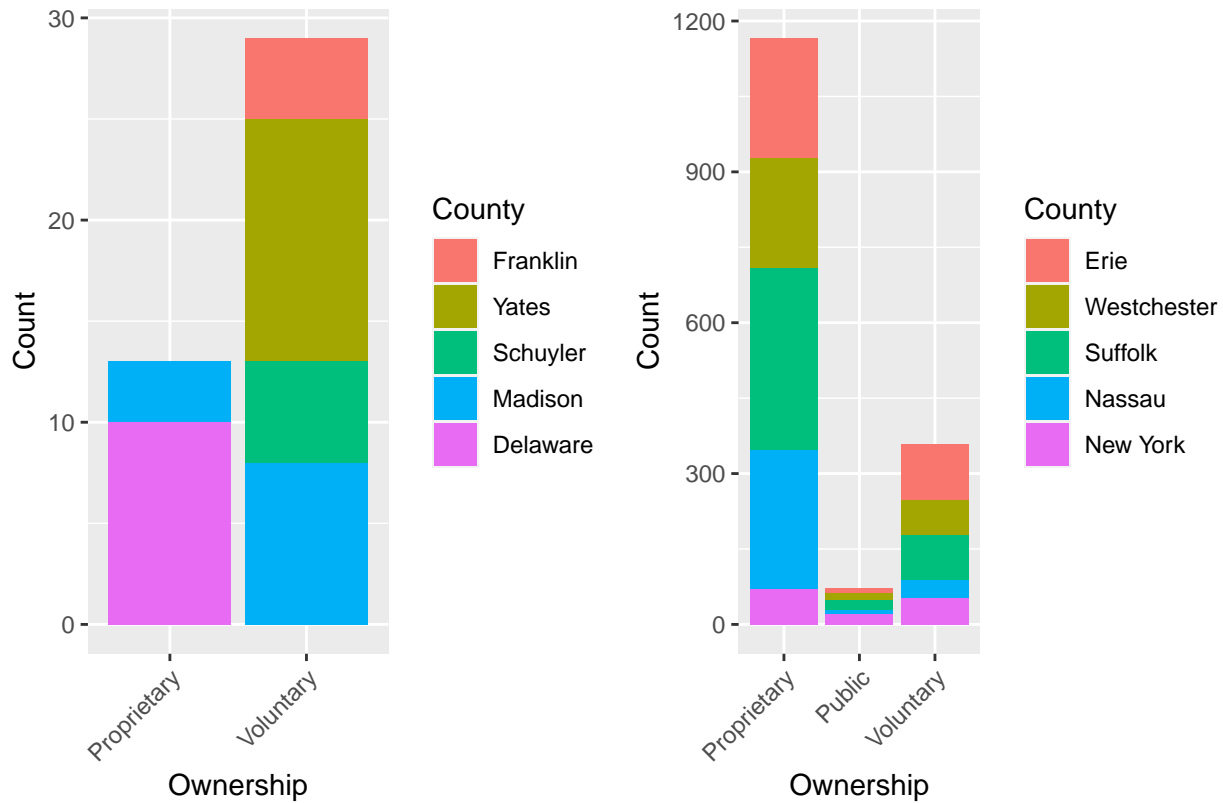
For nursing homes, there are two different key ownership types whether it be voluntary or proprietary. Voluntary is cheaper since it is not funded by either private shareholders and is in the nonprofit and community sector. Proprietary, on the other hand, is privately owned. We can conclude that privately owned nursing homes in New Yok State have seen lower fatality rates in residents due to COVID-19 than their voluntary counterparts.
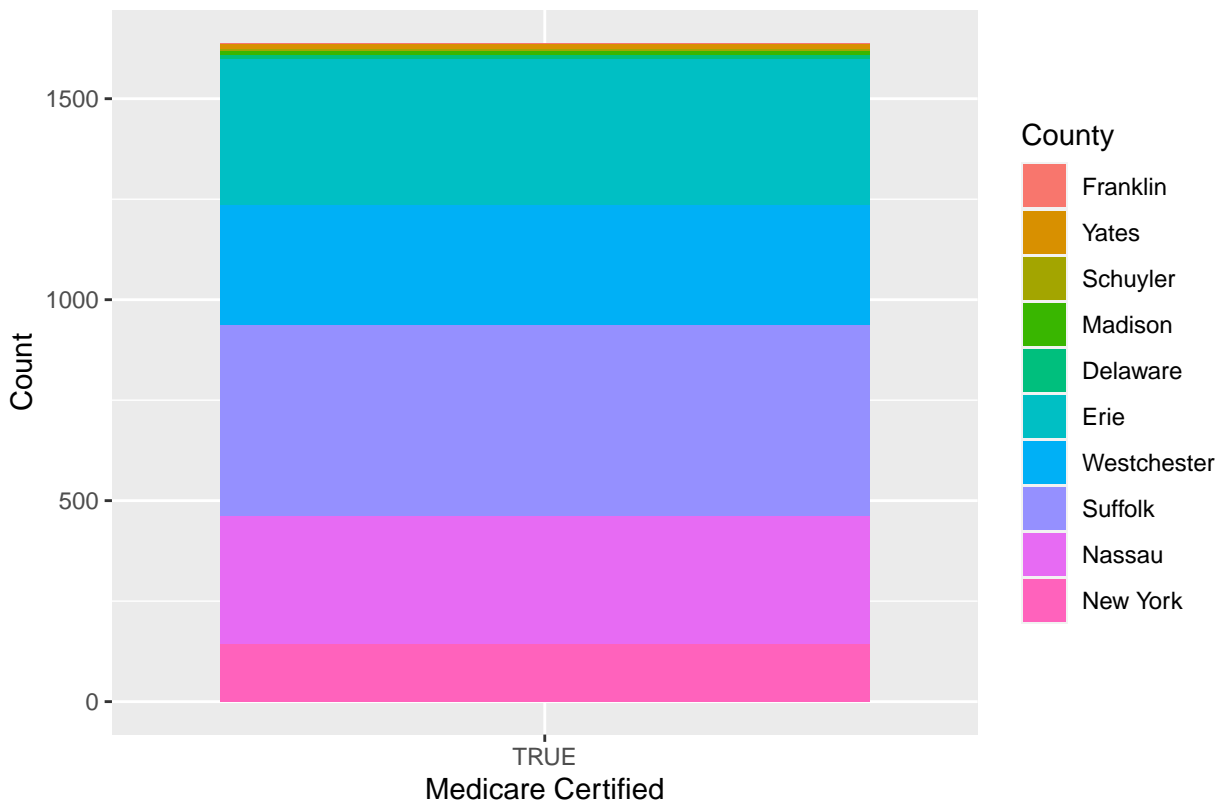
Supplemental Figures

Supplemental Figure 2: Nursing home facility complaints per 100 residents by Residents
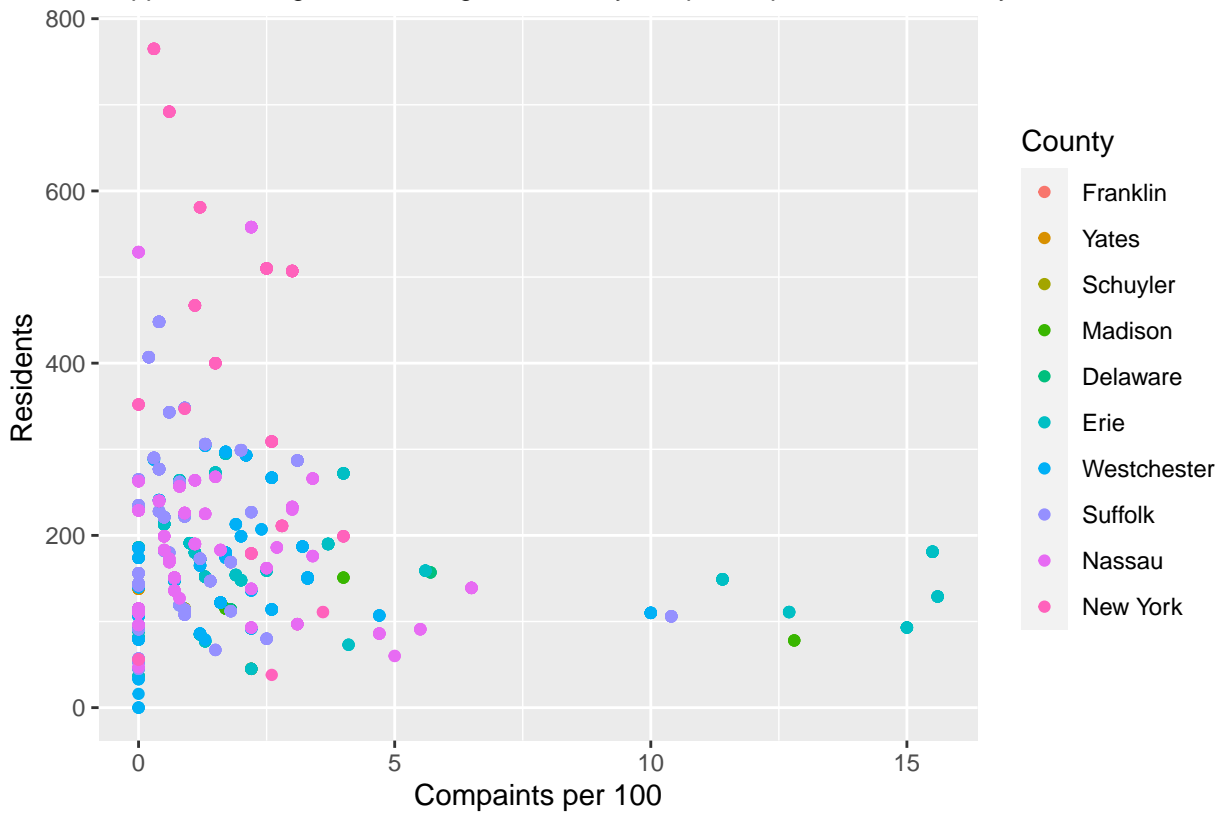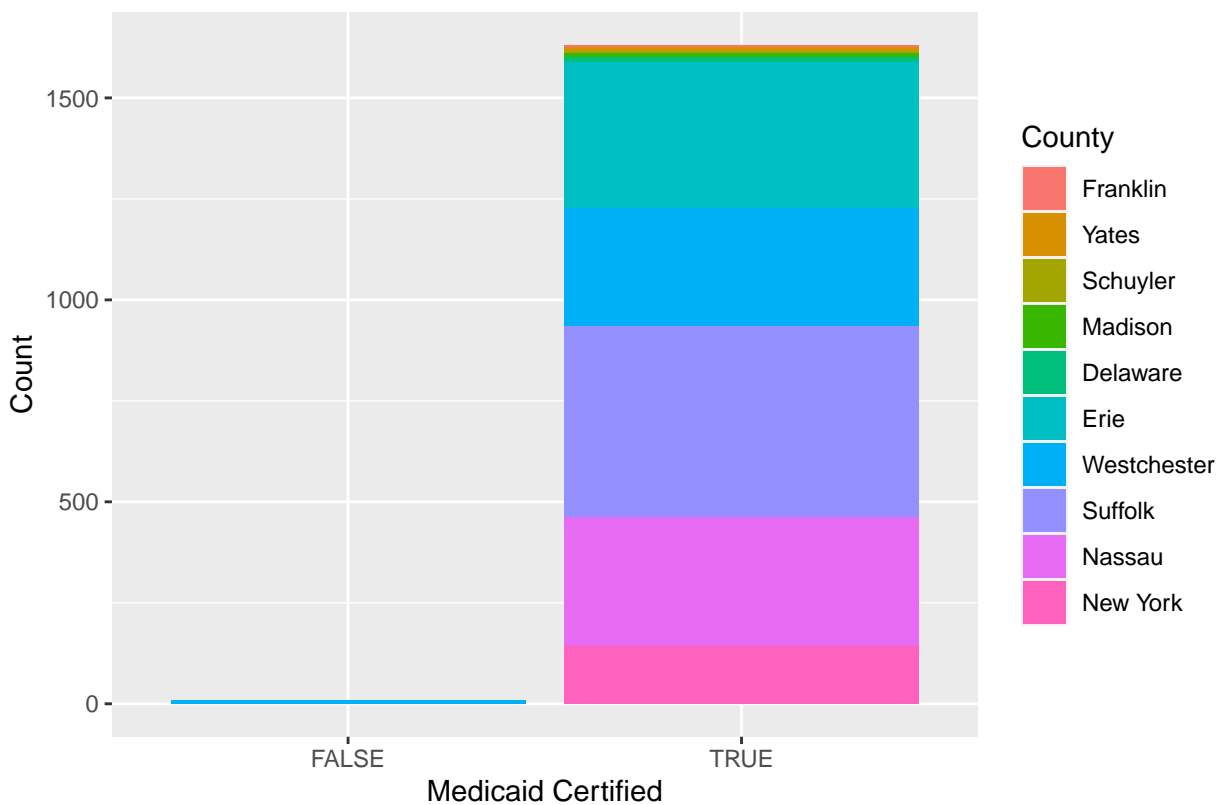

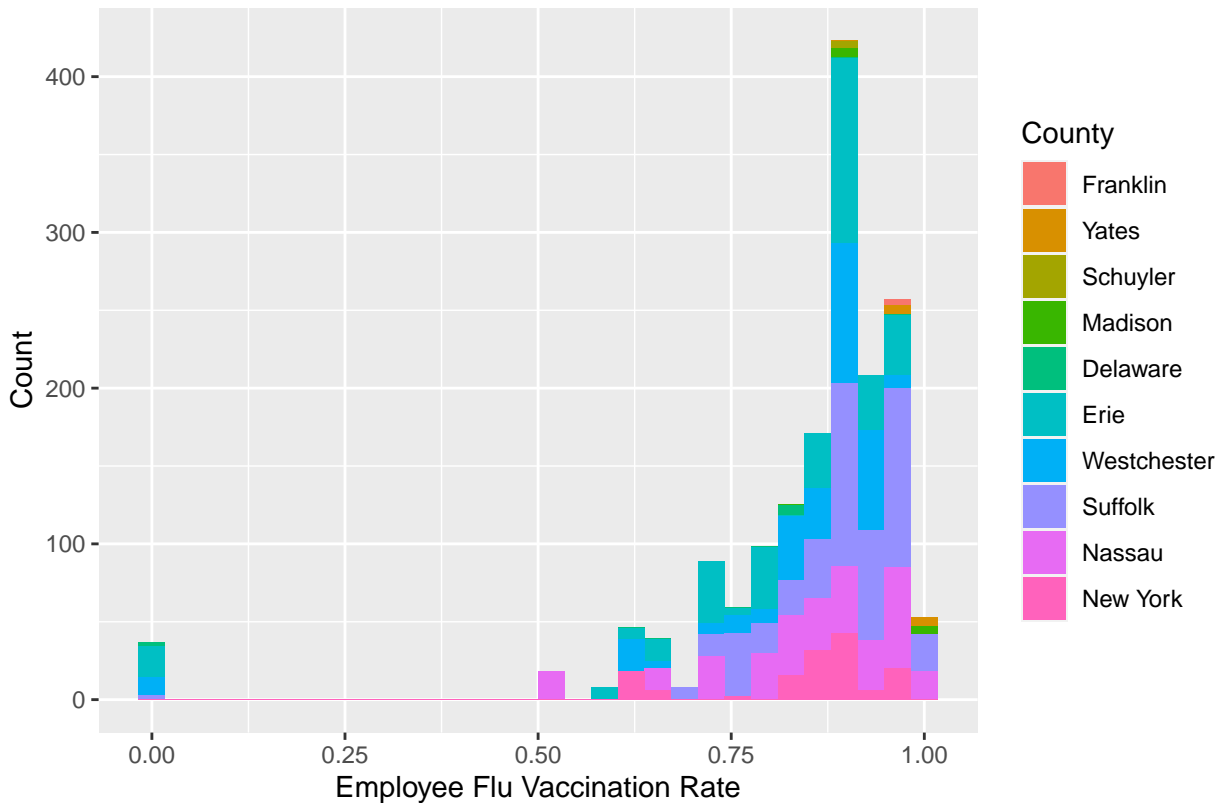Supplemental Figure 3: Nursing home facility total inspections by Residents
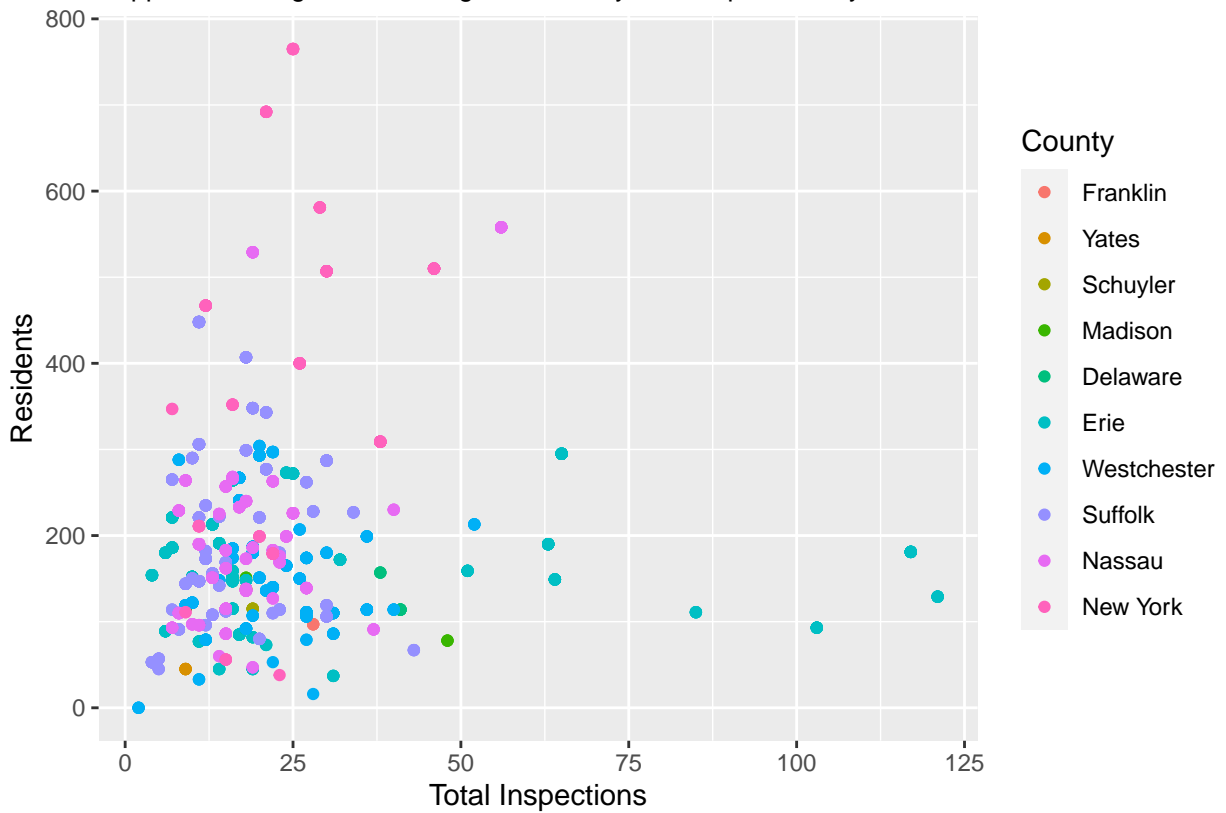
Supplemental Figure 4: Nursing home facility Medicare certification by county

Supplemental Figure 5: Nursing home facility Medicaid certification by county

Supplemental Figure 6: Nursing home facility employee flu vaccination rates by county

For my Supplemental Figures, we can see these data may not be the best variables to explain the discrepancy between nursing home fatalities. For example, in Supplemental Figures 5 and 6, we can see that virtually all nursing homes are Medicare and Medicaid certified, thus not a effective difference in nursing home quality. In Supplemental Figures 2 and 3 we see a slight trend in lower fatality nursing homes having more in house residents and and lesser complaints and inspections while some higher fatality nursing homes see the opposite, but to no extreme end. The majority of data points, low and high fatality percentages alike, cluster together rather than in two disparate groups. This, and the same clustering seen in Supplemental Figure 4, make for poor explanatory variables.

Summary

In this analysis, I observe the extent of the disparity in fatalities among New York State nursing homes by county and use facility information in order to determine which factors in nursing home quality effect said disparity. To achieve this, I calculated the percent of nursing home fatalities over the overall fatalities due to COVID-19 and analyzed the counties with the highest and lowest fatalities.

According to my findings, there are roughly two key factors that may have contributed to the disparity in nursing home fatalities due to COVID-19 across New York State. The most notable being the disparity in privately owned nursing homes in New York State have seen lower fatality rates in residents due to COVID-19 than their voluntarily owned nursing homes. This suggests that the amount of funding a nursing home receives may affect an older adult's chances of dying from COVID-19. While funding type acts as a surrogate for money funneled into a facility, if the data was available, I would look into the actual finances of these nursing homes to further develop this hypothesis.

The second factor is less reliable. Without population data, it's subjective to make claims that having more nursing homes would decrease an older adults chance of dying from COVID-19, since we can't look at the

ratio of nursing home adults to population density. If the data was available, I would take the proportion of adults over 65 per number of nursing homes and determine if we see similar trends between counties. For now, the data tentatively concludes that a county given more nursing homes is less likely to have a high count of older adults die due to COVID-19.

References

CDC. 2022. "COVID-19 Risks and Vaccine Information for Older Adults." U.S. Department of Health & Human Services: Centers for Disease Control; Prevention. https://www.cdc.gov/aging/covid19/covid19-older-adults.html.

Health, New York State Department of. 2022a. "Facility Info." New York State Department of Health: New York State Department of Health. https://health.data.ny.gov/Health/Nursing-Home-Profile/dypu-nabu.

———. 2022b. "New York State Statewide Covid19 Nursing Home and Adult Care Facility Fatalities." New York State Department of Health: New York State Department of Health. https://health.data.ny.gov/Health/New-York-State-Statewide-COVID-19-Nursing-Home-and/u2vg-th2g.

Times, The New York. 2022. "US County COVID-19 Data." 620 Eighth Avenue, New York, NY: The New York Times. https://github.com/nytimes/covid-19-data.