Gun Series

Ryan Larson

11/24/2020

Hospital Data

```
hosp_zcta <- read_csv("minnepop_1620_agg_zipfull.csv") %>%
  rename(zipcode = Zipcode) %>%
  arrange(zipcode, year, weekofyr) %>%
  select(-`_chk`)
```

ACS 5-Year Estimates

```
#adding in 5-year ACS data
census_api_key("ecda17575f4d914b502c70f2bae7a5f3d253792d")

year <- lst(2016, 2017, 2018, 2019)

acs <- map_dfr(
  year,
  ~ get_acs(geography = "zcta",
              variables = c("B01001_001E", "B03003_003E",
                            "B02001_003E", "B02001_002E",
                            "B02001_004E", "B02001_008E",
                            "B02001_005E", "B02001_006E",
                            "B02001_007E", "B11001_003E",
                            "B17001_002E", "B01002_001E",
                            "B09010_002E", "B06009_005E",
                            "B01001_002E", "B99233_005E"),
              output = "wide",
              survey = "acs5",
              year = .x),  .id = "year") %>%
  rename(total_pop = B01001_001E,
         white_pop = B02001_002E,
         black_pop = B02001_003E,
         na_pop = B02001_004E,
         asian_pop = B02001_005E,
         hpi_pop = B02001_006E,
         other_pop = B02001_007E,
         biracial_pop = B02001_008E,
         hisp_pop = B03003_003E,
         ssi_snap = B09010_002E, #snap, ssi, public cash transfers
```

```
         med_age = B01002_001E,
         mar_fam = B11001_003E,
         povlevel = B17001_002E,
         bach_degree = B06009_005E,
         male = B01001_002E,
         nowork_12 = B99233_005E) %>%
  select(-ends_with("M", ignore.case = F), -GEOID) %>%
  mutate(zcta = str_sub(NAME, 6)) %>%
  select(-NAME) %>%
  select(zcta, everything()) %>%
  mutate(year = as.numeric(year)) %>%
  mutate_at(vars(-zcta, -year, -total_pop, -med_age), list(~(./total_pop)*100))


#linear imputation of 2020 until 2020 ACS release (12/9/2021)
acs_2020 <- acs %>%
  complete(zcta, year = 2016:2020) %>%
  group_by(zcta) %>%
  mutate_at(vars(-zcta, -year),
            funs(if(sum(!is.na(.))<2) {.} else{na_interpolation(., option = "linear")})) %>%
  filter(year==2020)

acs_imp <- acs %>%
  rbind(acs_2020) %>%
  mutate(zcta = as.numeric(zcta))

#joining to hospital data
gun_panel <- hosp_zcta %>%
  left_join(acs_imp, by = c("zipcode"="zcta", "year"))


#SF geometries - get all MN ZCTAs
mn_zcta <- get_acs(geography = "zcta",
                   state = "MN",
                   variables = "B01001_001",
                   output = "wide",
                   year = 2019,
                   geometry = T,
                   survey = "acs5") %>%
  rename(zcta = GEOID,
         pop_2019 = B01001_001E) %>%
  select(-c(NAME, B01001_001M, pop_2019)) %>%
  mutate(zcta = as.numeric(zcta))
```

```
##   |                                                                              |
```

```
zcta_universe <- unique(hosp_zcta$zipcode)


#joining to panel
gun_panel <- gun_panel %>%
  left_join(mn_zcta, by = c("zipcode"="zcta"))
```

Aggregate Hospital Panel to Week-Level

```
#aggregated gun_panel to week-level
hosp_series <- gun_panel %>%
  group_by(year, weekofyr) %>%
  summarize(assault_tot = sum(assault_tot, na.rm = T),
            unintent_tot = sum(unintent_tot, na.rm = T),
            suicide_tot = sum(suicide_tot, na.rm = T),
            undeter_tot = sum(undeter_tot, na.rm = T),
            legal_tot = sum(legal_tot, na.rm = T),
            combined_tot = sum(combined_tot, na.rm = T),
            total_pop = sum(total_pop, na.rm = T)) %>%
  mutate(assault_incid_c = (assault_tot/total_pop)*1000,
         unintent_incid_c = (unintent_tot/total_pop)*1000,
         suicide_incid_c = (suicide_tot/total_pop)*1000,
         undeter_incid_c = (undeter_tot/total_pop)*1000,
         legal_incid_c = (legal_tot/total_pop)*1000,
         combined_incid_c = (combined_tot/total_pop)*1000) %>%
  ungroup() %>%
  mutate(week_id = row_number())
```

Police Data Week-Level

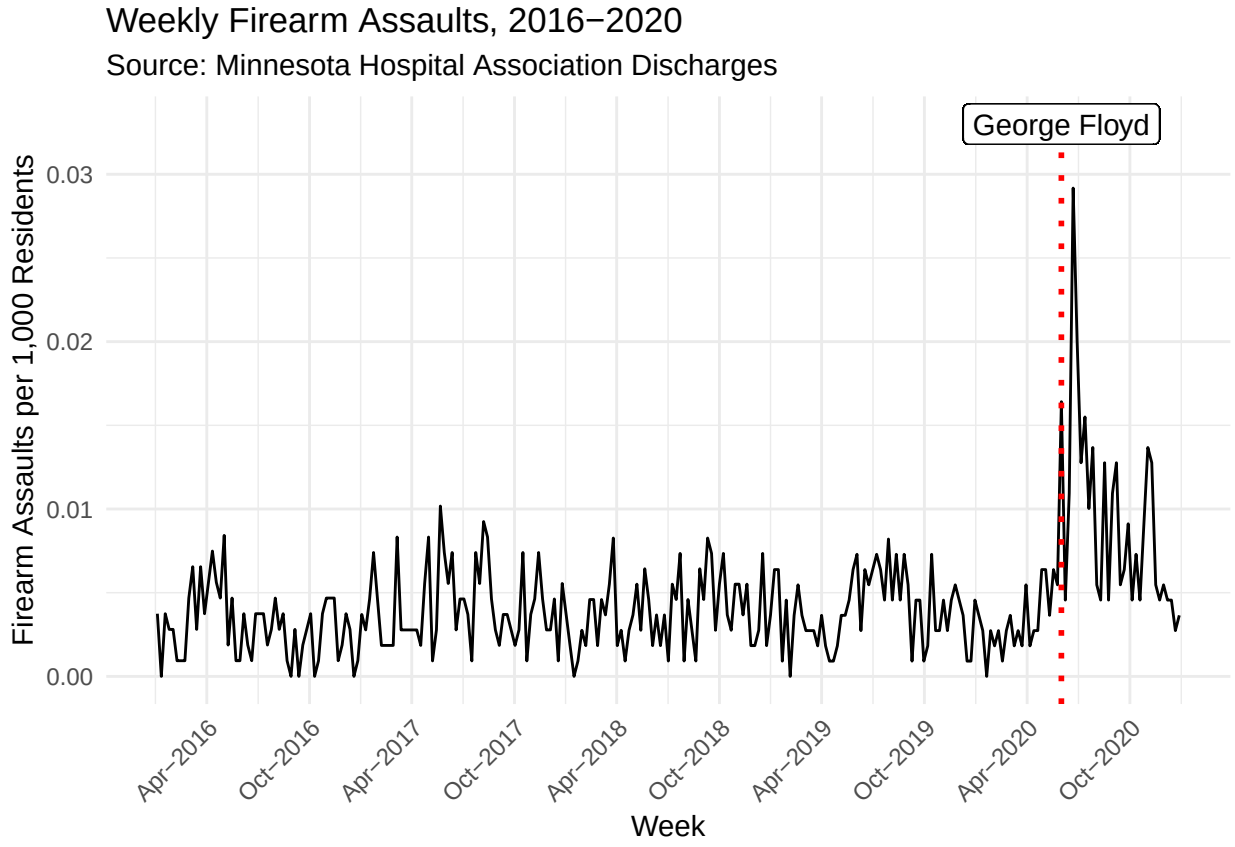
```
#Minneapolis POlice Department - Use of Force Dahsboard
uof <- read_csv("Police_Use_Of_Force.csv") %>%
  mutate(date=ymd_hms(ResponseDate),
         year=year(date),
         week=isoweek(date)) %>%
  group_by(year, week, .drop=F) %>%
  tally(name = "use_of_force") %>%
  arrange(year, week) %>%
  ungroup() %>%
  select(year, week, everything())

#merge onto series
series <- hosp_series %>%
  left_join(uof, by=c("year", "weekofyr"="week")) %>%
  mutate(use_of_force_rate = (use_of_force/total_pop)*1000)

#MPD Officer Involved Shootings
ois <- read_csv("Police_Officer_Involved_Shootings.csv") %>%
  mutate(date=ymd_hms(IncidentDate),
         year=year(date),
         week=isoweek(date)) %>%
  group_by(year, week, .drop=F) %>%
  tally(name = "off_inv_shooting") %>%
  arrange(year, week) %>%
  ungroup() %>%
  select(year, week, everything())

#merge onto series
```

```
series <- series %>%
  left_join(ois, by=c("year", "weekofyr"="week")) %>%
  mutate(off_inv_shooting_rate = (off_inv_shooting/total_pop)*1000)


#Minneapolis Police Department - Police Stops Dashboard
stop <- read_csv("Police_Stop_Data.csv") %>%
  mutate(date=ymd_hms(responseDate),
         year=year(date),
         week=isoweek(date)) %>%
  group_by(year, week, .drop=F) %>%
  tally(name = "police_stops")

#merge onto series
series <- series %>%
  left_join(stop, by = c("year", "weekofyr"="week")) %>%
  mutate(police_stop_rate = (police_stops/total_pop)*1000)

#Apple Mobility Data - don't have 2019, not using for now
apple_mpls <- apple_mobility %>%
  filter(region=="Minneapolis") %>%
  mutate(week = isoweek(date),
         year = year(date)) %>%
  group_by(year, week, .drop=F) %>%
  summarize(apple_mobility = mean(score, na.rm = T))

#New York Times COVID Case/Mortality Data
covid_hennepin <- nytcovcounty %>%
  mutate(week = isoweek(date),
         year = year(date)) %>%
  filter(county=="Hennepin" & state=="Minnesota" & year >=2019) %>%
  group_by(year, week, .drop=F) %>%
  summarize(covid_cases = sum(cases, na.rm = T),
            covid_deaths = sum(deaths, na.rm = T))

#filling 0s for pre-covid series
series <- series %>%
  left_join(covid_hennepin, by = c("year", "weekofyr"="week")) %>%
  mutate_at(vars(c(covid_cases, covid_deaths)), ~ifelse(is.na(.), 0, .))

#creating date variable
series <- series %>%
  mutate(begin_date = as.Date(paste(year, weekofyr, 1,sep = "-"), "%Y-%U-%u"),
         end_date = begin_date+weeks(1))
```

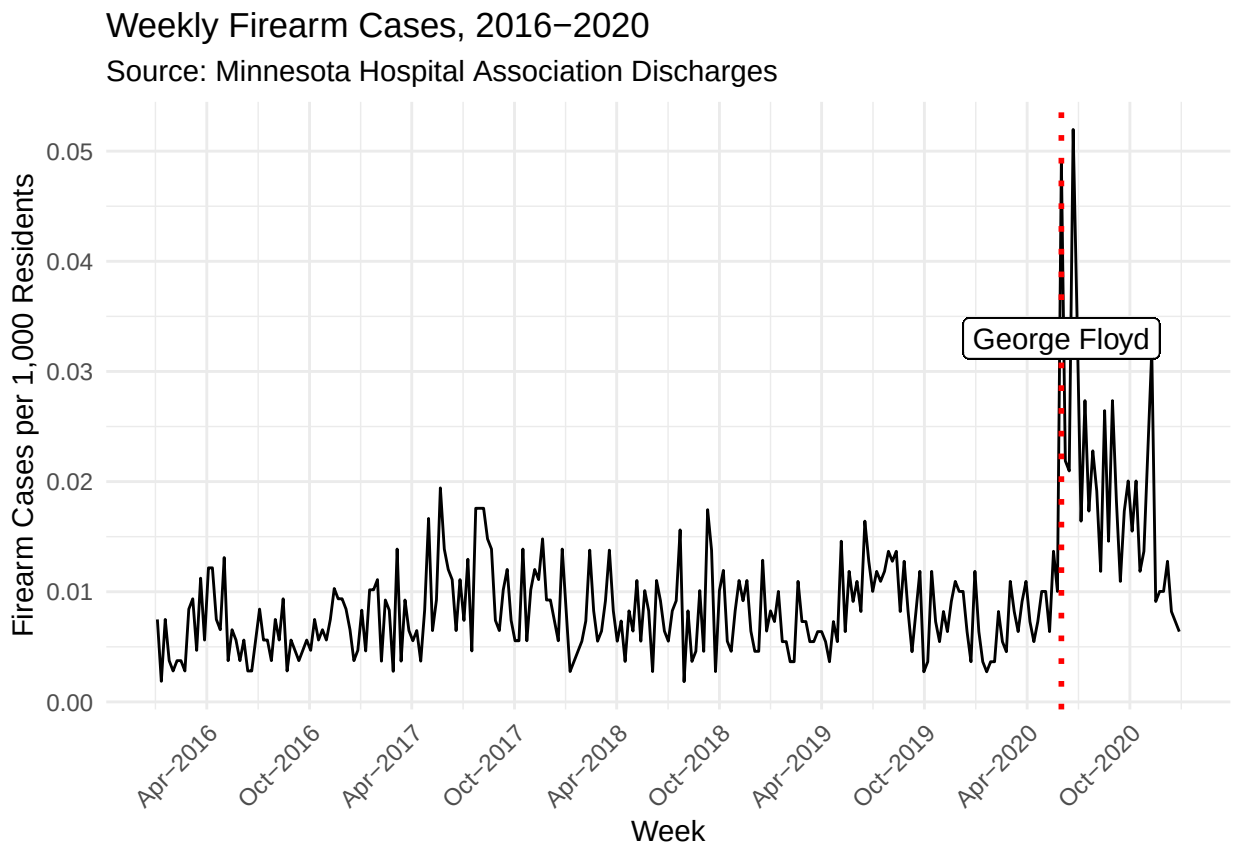
Time Series Vizualization

```
ggplot(series)+
  geom_line(aes(x=begin_date, y=assault_incid_c))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
```

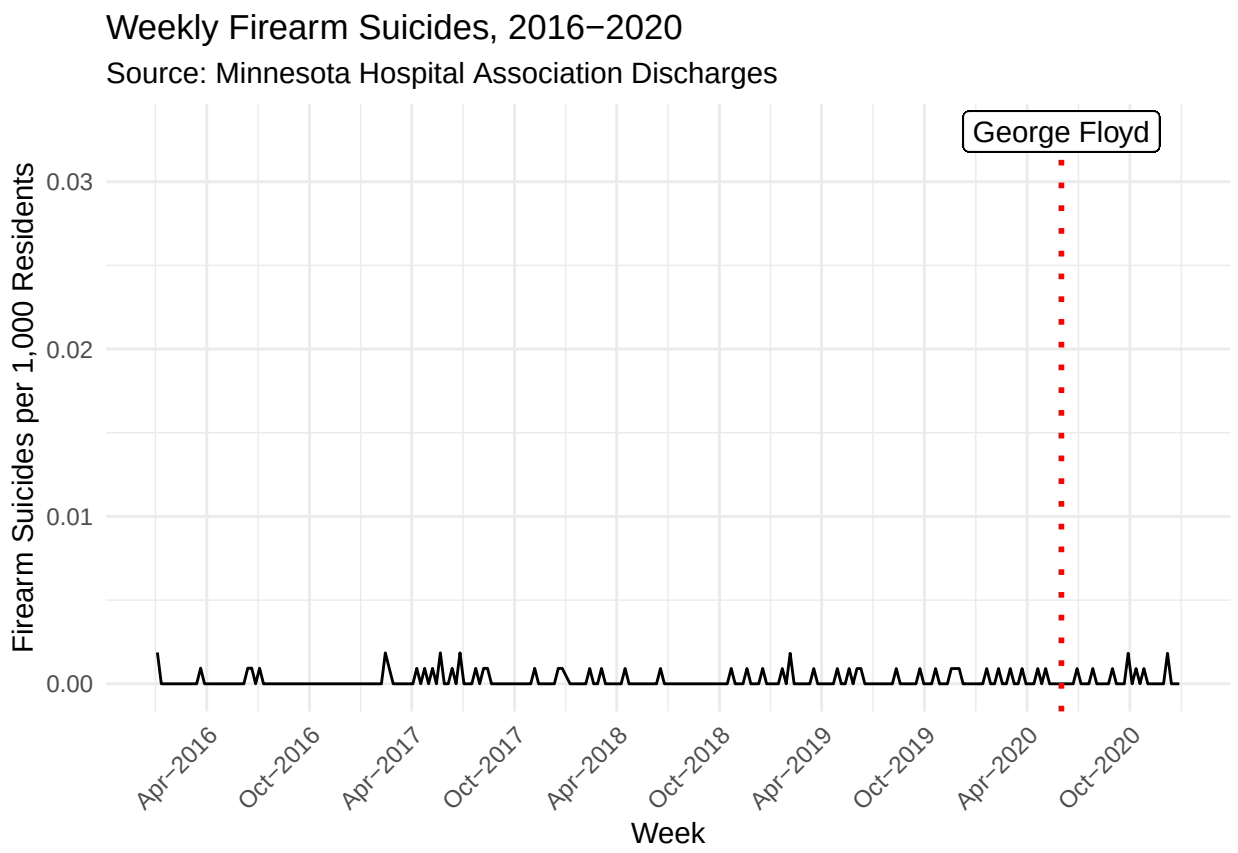
```
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Firearm Assaults, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association Discharges",
       x = "Week",
       y = "Firearm Assaults per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```




```
ggplot(series)+
  geom_line(aes(x=begin_date, y=combined_incid_c))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Firearm Cases, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association Discharges",
       x = "Week",
       y = "Firearm Cases per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```

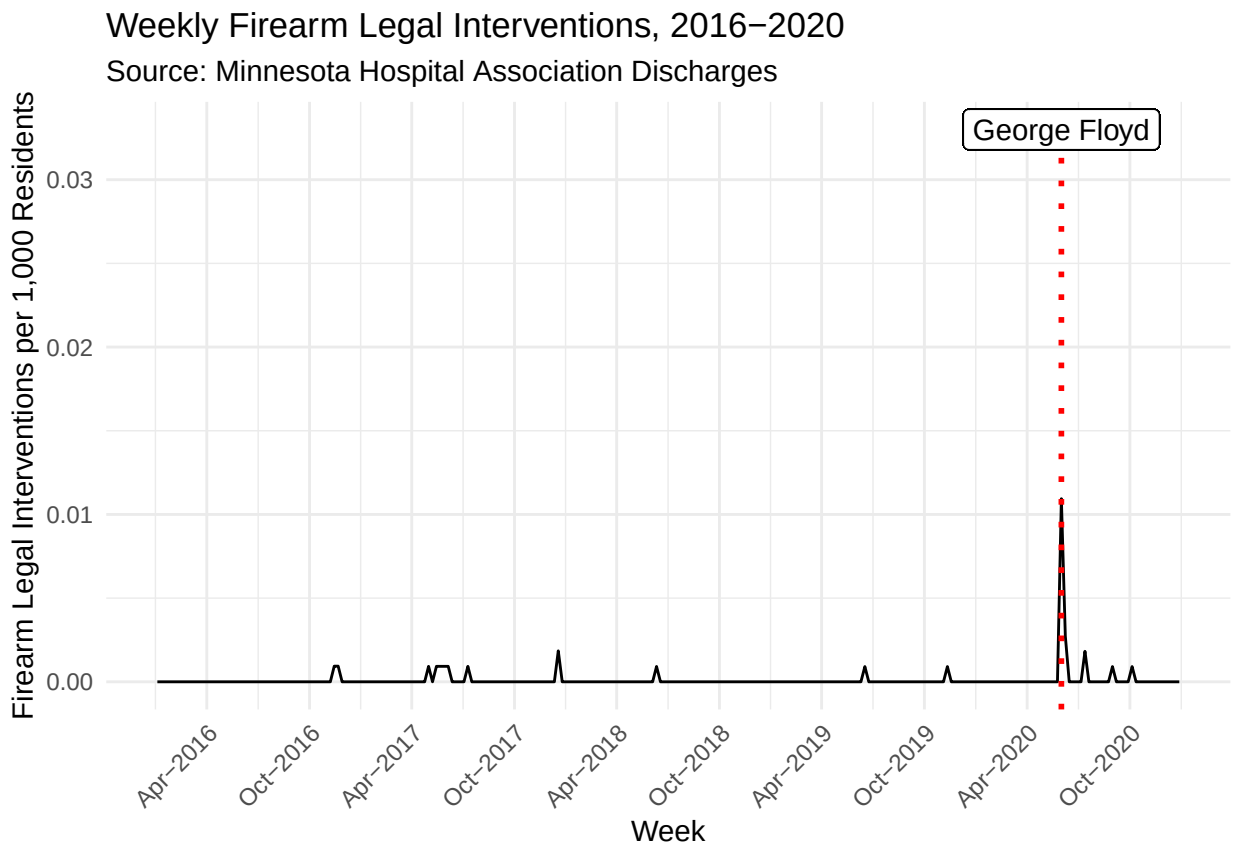



```
ggplot(series)+
  geom_line(aes(x=begin_date, y=suicide_incid_c))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Firearm Suicides, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association Discharges",
       x = "Week",
       y = "Firearm Suicides per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```

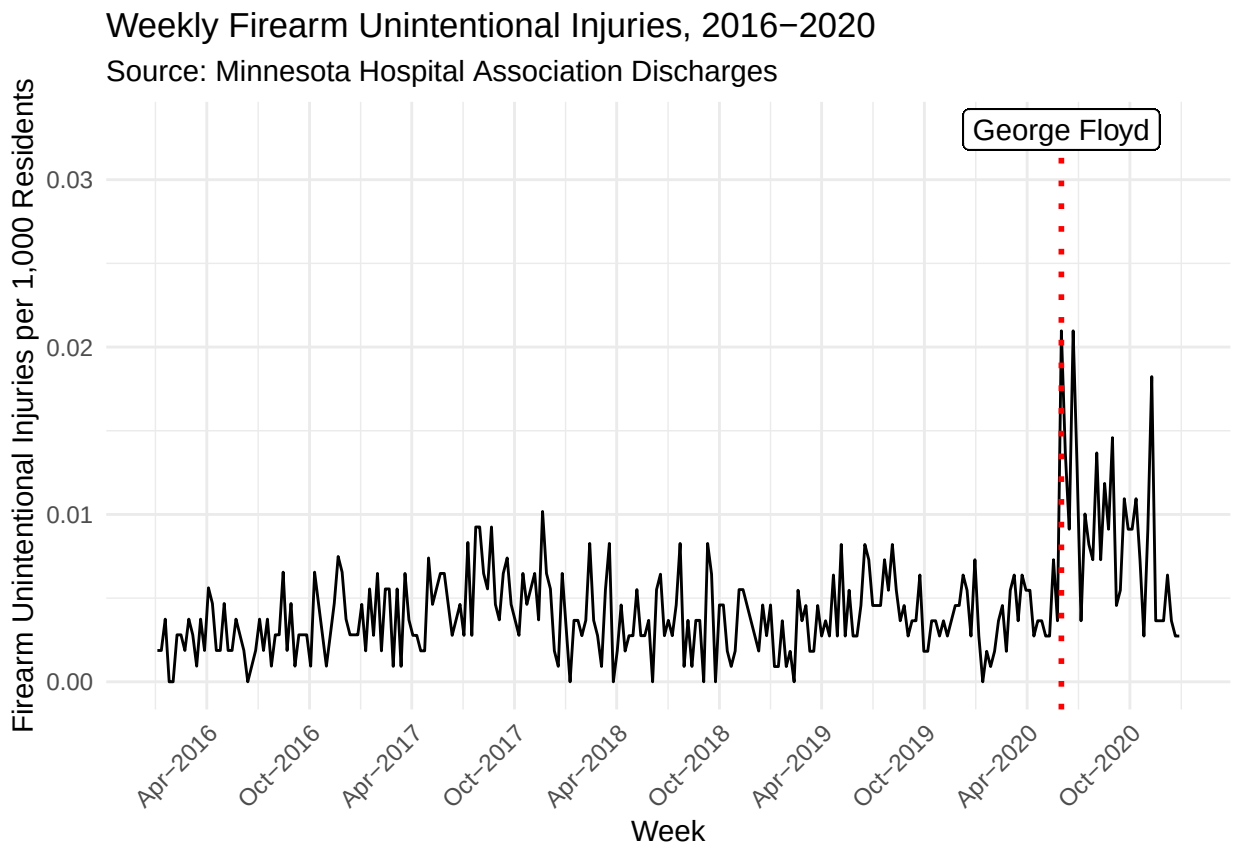



```
#are these so low because they don't end up in hospital?
```

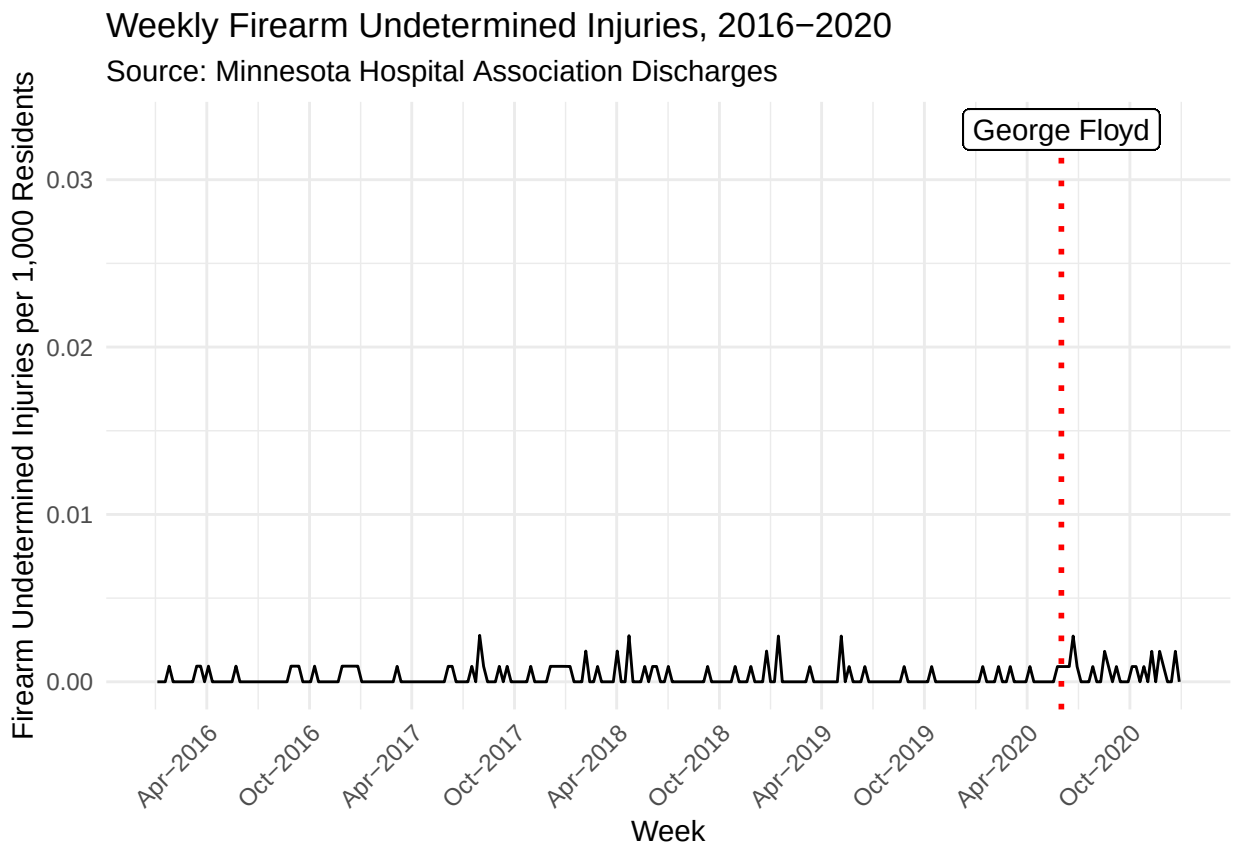
```
ggplot(series)+
  geom_line(aes(x=begin_date, y=legal_incid_c))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Firearm Legal Interventions, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association Discharges",
       x = "Week",
       y = "Firearm Legal Interventions per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```



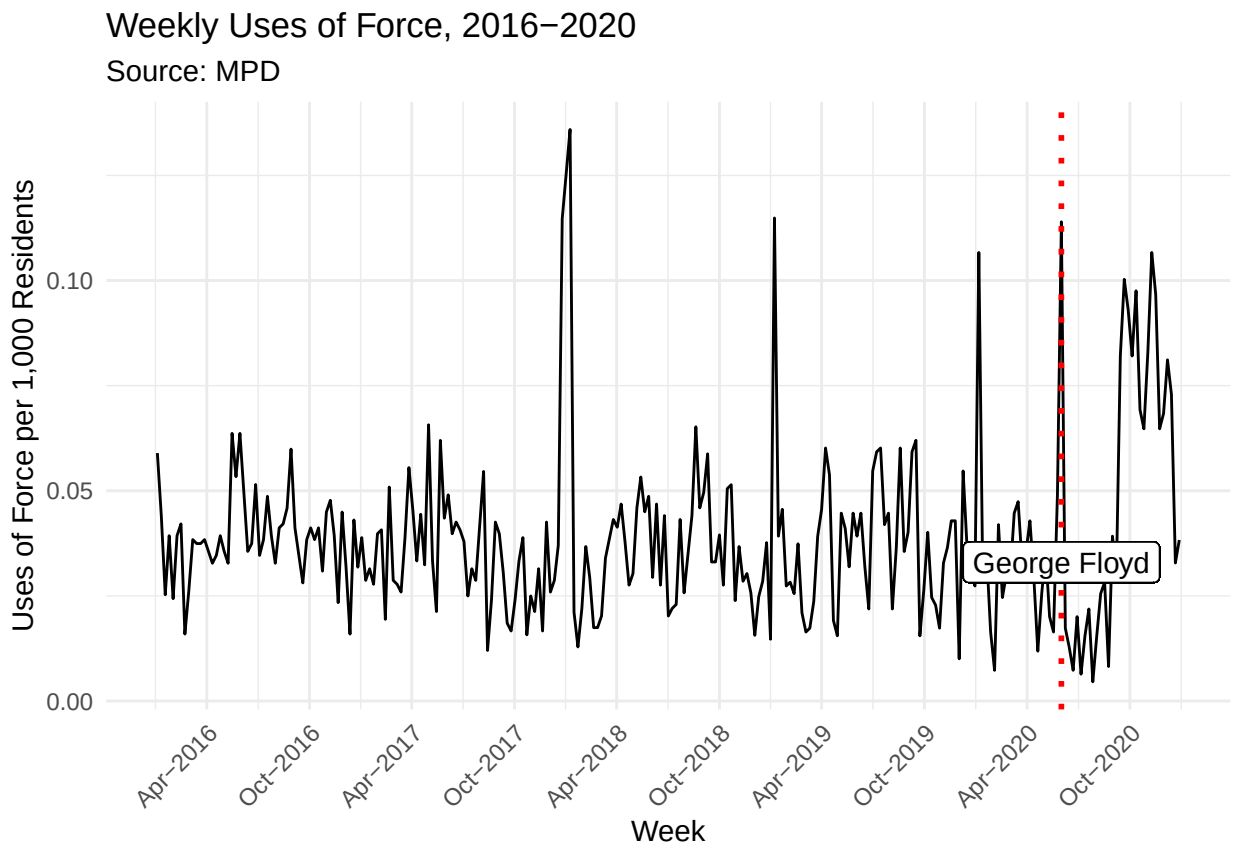

```
ggplot(series)+
  geom_line(aes(x=begin_date, y=unintent_incid_c))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Firearm Unintentional Injuries, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association Discharges",
       x = "Week",
       y = "Firearm Unintentional Injuries per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```



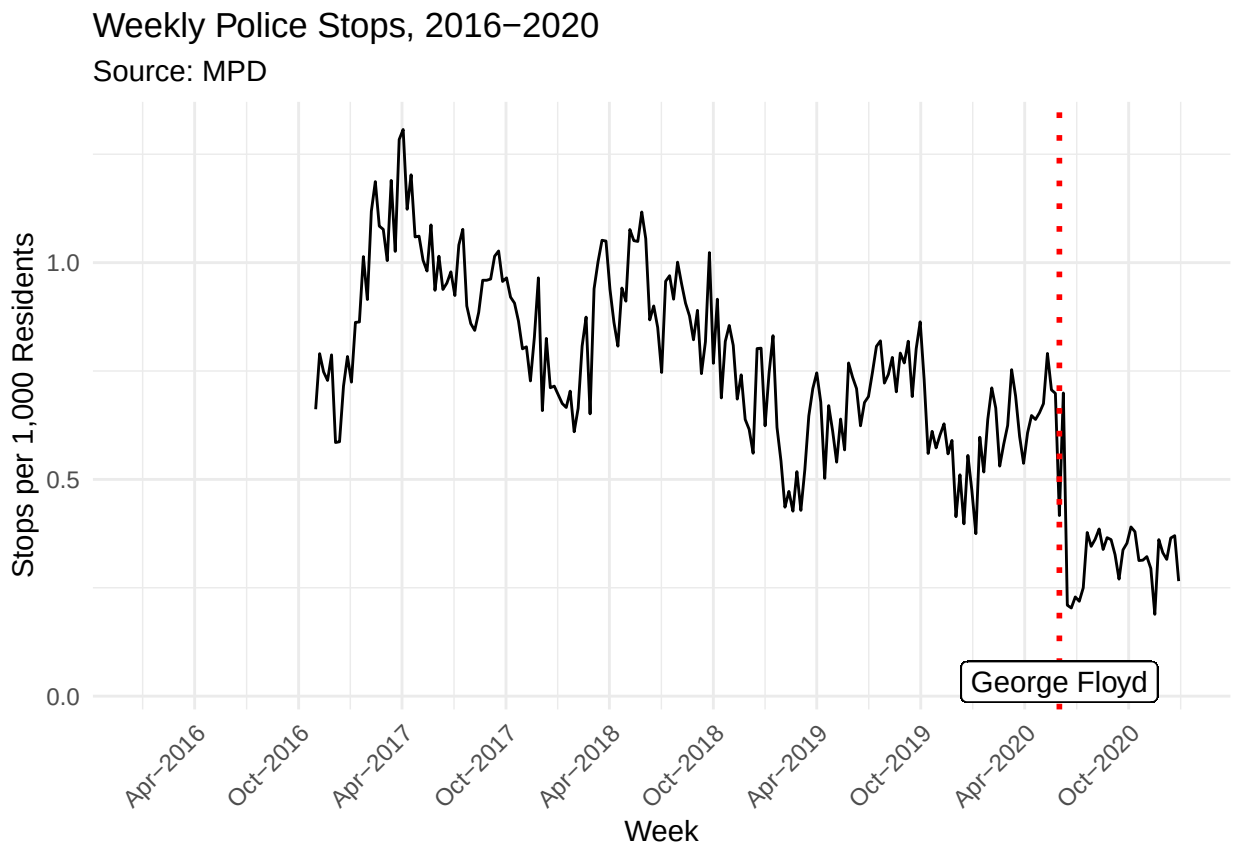

```
ggplot(series)+
  geom_line(aes(x=begin_date, y=undeter_incid_c))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Firearm Undetermined Injuries, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association Discharges",
       x = "Week",
       y = "Firearm Undetermined Injuries per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```




```
ggplot(series)+
  geom_line(aes(x=begin_date, y=use_of_force_rate))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Uses of Force, 2016-2020",
       subtitle = "Source: MPD",
       x = "Week",
       y = "Uses of Force per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```

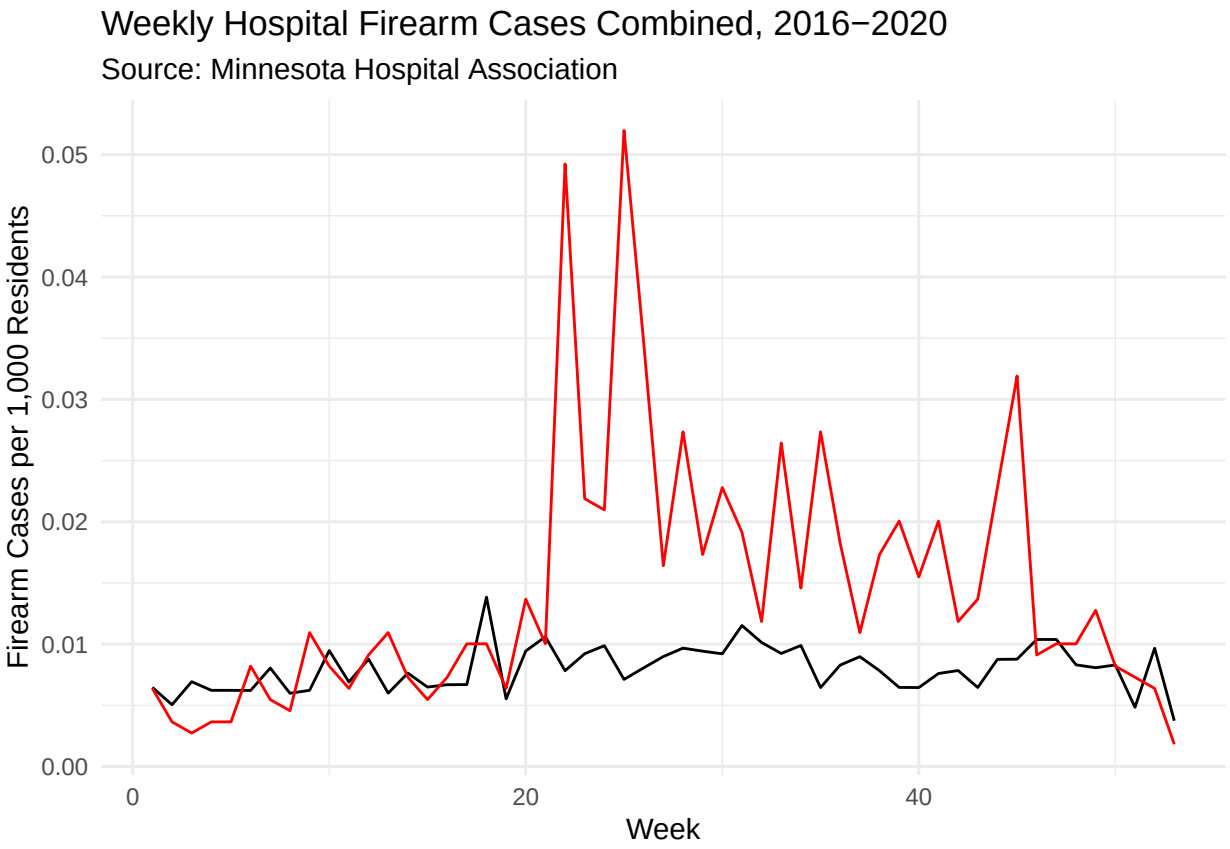



```
ggplot(series)+
  geom_line(aes(x=begin_date, y=police_stop_rate))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "6 months")+
  geom_vline(xintercept=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25
             linetype="dotted", color="red", size=1)+
  geom_label(aes(x=series$begin_date[series$year==2020 & series$weekofyr==isoweek(date("2020-05-25"))],
                 y=0.033),
             label = "George Floyd", show.legend = FALSE)+
  labs(title = "Weekly Police Stops, 2016-2020",
       subtitle = "Source: MPD",
       x = "Week",
       y = "Stops per 1,000 Residents")+
  theme_minimal()+
  theme(axis.text.x=element_text(angle=45, hjust=1))
```




```
#need ois plot here; fix missings
```

```
pre <- series %>%
  filter(year <= 2019) %>%
  group_by(weekofyr) %>%
  summarize(combined = mean(combined_incid_c, na.rm = T))

post <- series %>%
  filter(year >= 2020)

ggplot()+
  geom_line(data = pre, aes(x=weekofyr, y=combined))+
  geom_line(data = post, aes(x=weekofyr, y=combined_incid_c), color = "red")+
   labs(title = "Weekly Hospital Firearm Cases Combined, 2016-2020",
       subtitle = "Source: Minnesota Hospital Association ",
       x = "Week",
       y = "Firearm Cases per 1,000 Residents")+
  theme_minimal()
```



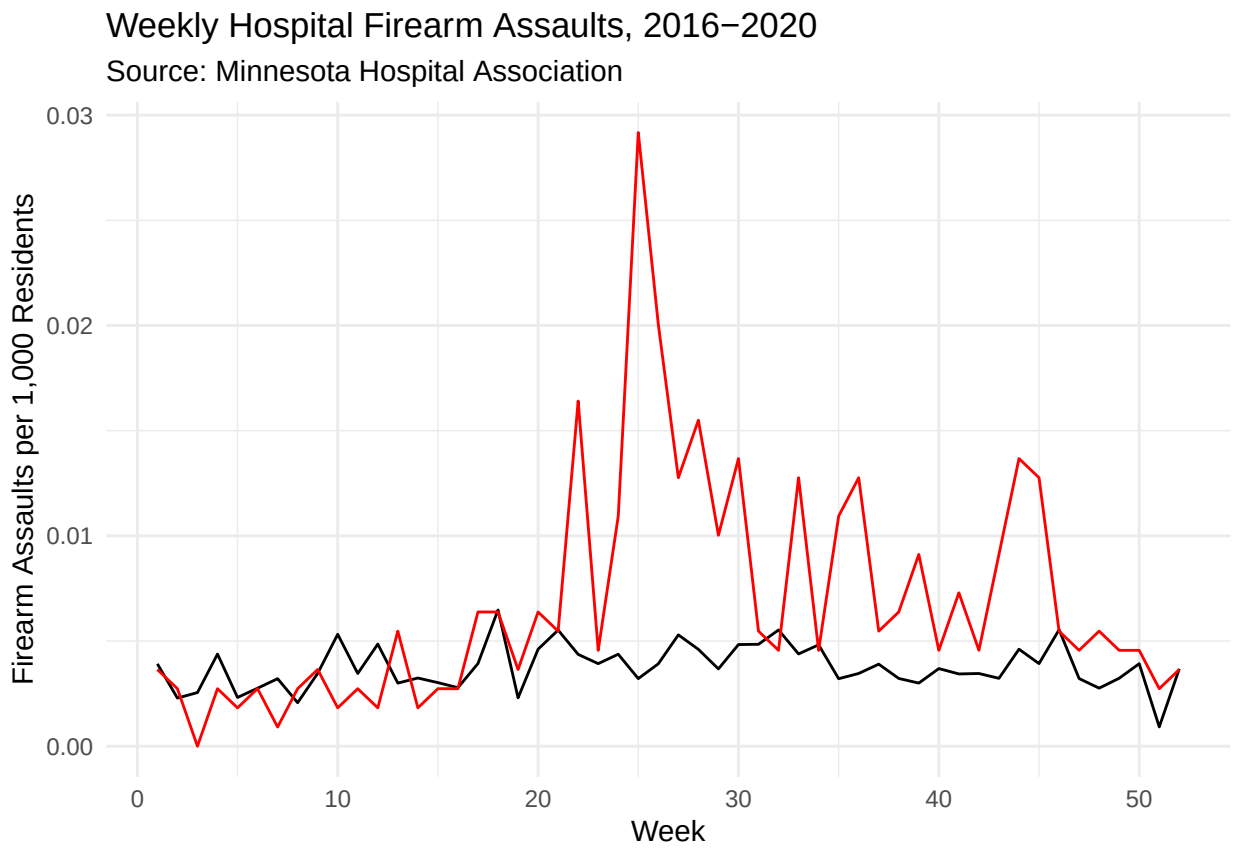
Time Series Analysis

```
window_df <- function(df, outcome, end.date, start.date, end.pre.date) {

  ### Function to generate windowed data frame objects
  ###   for a full time period (start through end)
  ###   and for a pre-treatment period (start through pre)
  ### Inputs
  ###  df: data frame to be windowed
  ###  outcome: column number of a crime or arrest category (integer)
  ###  end.date: last date for full time period (date format)
  ###  start.date: first date for full time period (date format)
  ###  end.pre.date: last date for pre-treatment period (date format)
  ### Returns:
  ###  list containing 2 data frame objects

  ## select variables from data frame and filter

  df.windowed.pre <- select(df, begin_date, year, weekofyr, y = outcome) %>%
                     dplyr::filter(begin_date >= start.date,
                            begin_date <= end.pre.date)
  df.windowed.post <- select(df, begin_date, year, weekofyr, y = outcome) %>%
                      dplyr::filter(begin_date >= end.pre.date)
  df.windowed.all <- select(df, begin_date, year, weekofyr, y = outcome) %>%
```

```
                    dplyr::filter(begin_date >= start.date,
                           begin_date <= end.date)
  df.pre.agg <- df.windowed.pre %>%
    group_by(weekofyr) %>%
    summarize(y = mean(y, na.rm = T))


  ## return list

  list.df <- list(df.windowed.pre, df.windowed.post, df.windowed.all, df.pre.agg)

  return(list.df)
}

pre_2020 <- window_df(series,
               outcome = "assault_incid_c",
               end.date = "2020-12-31",
               start.date = "2016-01-01",
               end.pre.date = "2020-01-01")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use ‘all_of(outcome)‘ instead of ‘outcome‘ to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
ggplot()+
  geom_line(data = as.data.frame(pre_2020[[4]]), aes(x=weekofyr, y=y))+
  geom_line(data = as.data.frame(pre_2020[[2]]), aes(x=weekofyr, y=y), color = "red")+
   labs(title = "Weekly Hospital Firearm Assaults, 2016-2020",
        subtitle = "Source: Minnesota Hospital Association ",
        x = "Week",
        y = "Firearm Assaults per 1,000 Residents")+
  theme_minimal()
```




```
#models add in
```

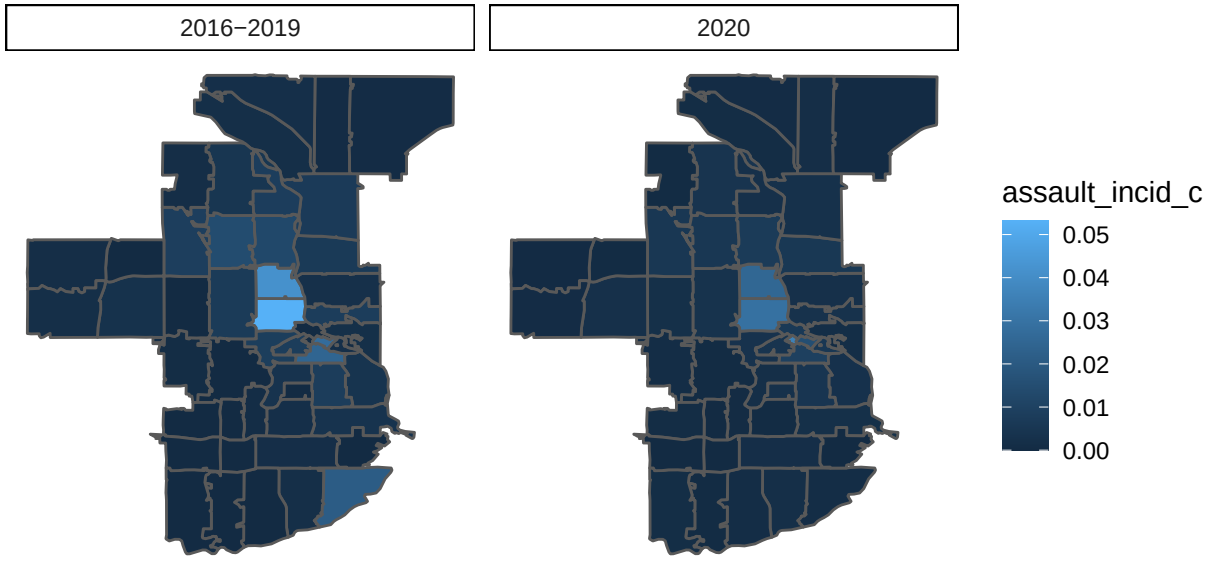
ZCTA-Week Level Police Data

```
#Minneapolis POlice Department - Use of Force Dahsboard
uof_spatial <- read_csv("Police_Use_Of_Force.csv")  %>%
  mutate(date=ymd_hms(ResponseDate),
         year=year(date),
         week=isoweek(date)) %>%
  select(OBJECTID, year, week, X, Y, Race) %>%
  st_as_sf(coords = c("X", "Y"), crs = "NAD83", remove=F) %>%
  mutate(intersection = as.integer(st_intersects(geometry, mn_zcta)),
         zcta = ifelse(is.na(intersection), NA, mn_zcta$zcta[intersection])) %>%
  st_drop_geometry() %>%
  filter(!is.na(zcta) & year >= 2016) %>%
  group_by(year, week, zcta, Race, .drop=F) %>%
  tally(name = "use_of_force") %>%
  filter(!is.na(Race) & Race!="not recorded") %>%
  ungroup() %>%
  complete(year, week, zcta=zcta_universe, Race, fill = list(use_of_force = 0)) %>%
  arrange(year, week, zcta, Race) %>%
  mutate(race = str_to_lower(Race)) %>%
  select(-Race) %>%
  pivot_wider(names_from = race,
```

```
              values_from = use_of_force,
              values_fill = 0,
              names_glue = "{race}_{.value}") %>%
  mutate(total_use_of_force = asian_use_of_force+black_use_of_force+`native american_use_of_force`+
           `other / mixed race_use_of_force`+`pacific islander_use_of_force`+unknown_use_of_force+
            white_use_of_force)

#MPD Stop Dashboard
stop_spatial <- read_csv("Police_Stop_Data.csv") %>%
  mutate(date=ymd_hms(responseDate),
         year=year(date),
         week=isoweek(date)) %>%
  select(OBJECTID, year, week, lat, long, race) %>%
  st_as_sf(coords = c("long", "lat"), crs = "NAD83", remove=F) %>%
  mutate(intersection = as.integer(st_intersects(geometry, mn_zcta)),
         zcta = ifelse(is.na(intersection), NA, mn_zcta$zcta[intersection])) %>%
  st_drop_geometry() %>%
  filter(!is.na(zcta) & year >= 2016 & !zcta %in% c(55111,55113,55114)) %>%
  group_by(year, week, zcta, race, .drop=F) %>%
  tally(name = "police_stops") %>%
   filter(!is.na(race) & race!="not recorded") %>%
  ungroup() %>%
  complete(year, week, zcta=zcta_universe, race, fill = list(police_stops = 0)) %>%
  mutate(race = str_to_lower(race)) %>%
  arrange(year, week, zcta, race) %>%
  pivot_wider(names_from = race,
              values_from = police_stops,
              values_fill = 0,
              names_glue = "{race}_{.value}") %>%
  mutate(total_police_stops = asian_police_stops+black_police_stops+
         `east african_police_stops`+latino_police_stops+`native american_police_stops`+
           other_police_stops+unknown_police_stops+white_police_stops)

ois_spatial <- read_csv("Police_Officer_Involved_Shootings.csv") %>%
  mutate(date=ymd_hms(IncidentDate),
         year=year(date),
         week=isoweek(date)) %>%
  select(OBJECTID, year, week, CenterLatitude, CenterLongitude, SubjectOfForceRace) %>%
  rename(race = SubjectOfForceRace,
         lat = CenterLatitude,
         long = CenterLongitude) %>%
  st_as_sf(coords = c("long", "lat"), crs = "NAD83", remove=F) %>%
  mutate(intersection = as.integer(st_intersects(geometry, mn_zcta)),
         zcta = ifelse(is.na(intersection), NA, mn_zcta$zcta[intersection])) %>%
  st_drop_geometry() %>%
  filter(!is.na(zcta) & year >= 2016) %>%
  group_by(year, week, zcta, race, .drop=F) %>%
  tally(name = "police_shootings") %>%
  filter(!is.na(race) & race!="not recorded") %>%
  ungroup() %>%
  complete(year=2016:2021, week=1:53, zcta=zcta_universe, race, fill = list(police_shootings = 0)) %>%
  mutate(race = str_to_lower(race)) %>%
  arrange(year, week, zcta, race) %>%
```

```
  pivot_wider(names_from = race,
              values_from = police_shootings,
              values_fill = 0,
              names_glue = "{race}_{.value}") %>%
  mutate(total_police_shootings = asian_police_shootings+black_police_shootings+
         hispanic_police_shootings+other_police_shootings+
           unknown_police_shootings+white_police_shootings)

gun_panel <- gun_panel %>%
  left_join(uof_spatial, by = c("year", "weekofyr"="week", "zipcode"="zcta")) %>%
  left_join(stop_spatial, by = c("year", "weekofyr"="week", "zipcode"="zcta")) %>%
  left_join(ois_spatial, by = c("year", "weekofyr"="week", "zipcode"="zcta"))
```

```
#aggregate to zip-level over years
zip_level <- gun_panel %>%
  mutate(post = ifelse(year >= 2020, "2016-2019", "2020")) %>%
  group_by(zipcode, post) %>%
  summarize(assault_tot = sum(assault_tot, na.rm = T),
            unintent_tot = sum(unintent_tot, na.rm = T),
            suicide_tot = sum(suicide_tot, na.rm = T),
            undeter_tot = sum(undeter_tot, na.rm = T),
            legal_tot = sum(legal_tot, na.rm = T),
            combined_tot = sum(combined_tot, na.rm = T),
            total_pop = sum(total_pop, na.rm = T)) %>%
  mutate(assault_incid_c = (assault_tot/total_pop)*1000,
         unintent_incid_c = (unintent_tot/total_pop)*1000,
         suicide_incid_c = (suicide_tot/total_pop)*1000,
         undeter_incid_c = (undeter_tot/total_pop)*1000,
         legal_incid_c = (legal_tot/total_pop)*1000,
         combined_incid_c = (combined_tot/total_pop)*1000) %>%
  ungroup() %>%
  left_join(mn_zcta, by = c("zipcode"="zcta"))


ggplot(zip_level, aes(geometry = geometry, fill = assault_incid_c)) +
  geom_sf() +
  coord_sf(crs = "+proj=merc")+
  facet_wrap(~post)+
  theme(axis.text = element_blank(),
  axis.line = element_blank(),
  axis.ticks = element_blank(),
  panel.border = element_blank(),
  panel.grid = element_blank(),
  axis.title = element_blank(),
  panel.background = element_blank(),
  panel.grid.major = element_line(colour="transparent"),
  plot.subtitle = element_text(face="italic"),
  strip.background = element_rect(fill = "white",
                colour = "black"))
```




```
#may need to make sure all ZCTAs here are MPLS?
```