USFirearm

Syed Baryalay

2023-04-19

Generate some insightful visualizations to display this data.

Does the rate of change in total firearms background checks over time vary across states?

```
suppressMessages(library(tidyverse))
suppressMessages(library(dplyr))
suppressMessages(library(tidyr))
suppressMessages(library(forecast))
suppressMessages(library(ggplot2))

url <- "https://raw.githubusercontent.com/BuzzFeedNews/nics-firearm-background-checks/master/data/nics-f

firearm <- read.csv(url, header = TRUE)
glimpse(firearm)
```

```
## Rows: 16,115
## Columns: 27
## $ month                    <chr> "2023-03", "2023-03", "2023-03", "2023-03", ~
## $ state                    <chr> "Alabama", "Alaska", "Arizona", "Arkansas", ~
## $ permit                   <int> 14315, 354, 12965, 4460, 25878, 11466, 9850,~
## $ permit_recheck           <int> 263, 7, 1377, 595, 13200, 5, 640, 0, 0, 0, 0~
## $ handgun                  <int> 23132, 3295, 20940, 8827, 40714, 24432, 7098~
## $ long_gun                 <int> 15158, 2610, 10397, 6436, 26473, 15658, 2405~
## $ other                    <int> 1314, 397, 1772, 573, 5455, 2504, 1076, 94, ~
## $ multiple                 <int> 1269, 221, 1249, 547, 0, 2464, 0, 94, 8, 335~
## $ admin                    <int> 0, 0, 0, 4, 0, 0, 5, 0, 3, 0, 0, 0, 0, 0, 0,~
## $ prepawn_handgun          <int> 17, 1, 7, 14, 1, 0, 0, 0, 0, 16, 16, 0, 0, 3~
## $ prepawn_long_gun         <int> 3, 0, 2, 11, 1, 0, 0, 0, 0, 6, 9, 0, 0, 1, 0~
## $ prepawn_other            <int> 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,~
## $ redemption_handgun       <int> 2564, 136, 1446, 1280, 746, 0, 0, 30, 0, 405~
## $ redemption_long_gun      <int> 1116, 95, 503, 918, 409, 0, 0, 6, 0, 975, 88~
## $ redemption_other         <int> 16, 2, 5, 3, 16, 0, 0, 0, 0, 6, 13, 0, 0, 1,~
## $ returned_handgun         <int> 46, 29, 260, 0, 1731, 340, 0, 63, 2, 1495, 7~
## $ returned_long_gun        <int> 3, 15, 22, 0, 927, 46, 0, 0, 0, 138, 0, 0, 1~
## $ returned_other           <int> 0, 0, 0, 0, 82, 0, 0, 0, 54, 4, 0, 0, 0, 2, ~
## $ rentals_handgun          <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
## $ rentals_long_gun         <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
## $ private_sale_handgun     <int> 35, 0, 7, 7, 8581, 0, 627, 51, 0, 308, 12, 0~
## $ private_sale_long_gun    <int> 34, 1, 1, 12, 3392, 0, 196, 23, 0, 204, 3, 0~
## $ private_sale_other       <int> 8, 0, 2, 4, 626, 0, 87, 0, 0, 67, 1, 0, 0, 0~
## $ return_to_seller_handgun <int> 0, 1, 0, 0, 116, 0, 0, 0, 0, 58, 0, 0, 0, 1,~
## $ return_to_seller_long_gun <int> 0, 0, 1, 0, 51, 0, 0, 0, 0, 41, 0, 0, 0, 0, ~
```

```
## $ return_to_seller_other    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
## $ totals                    <int> 59294, 7164, 50957, 23692, 128399, 56915, 21~
```

Cleaning and filtering original data for analysis

```
library(dplyr)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

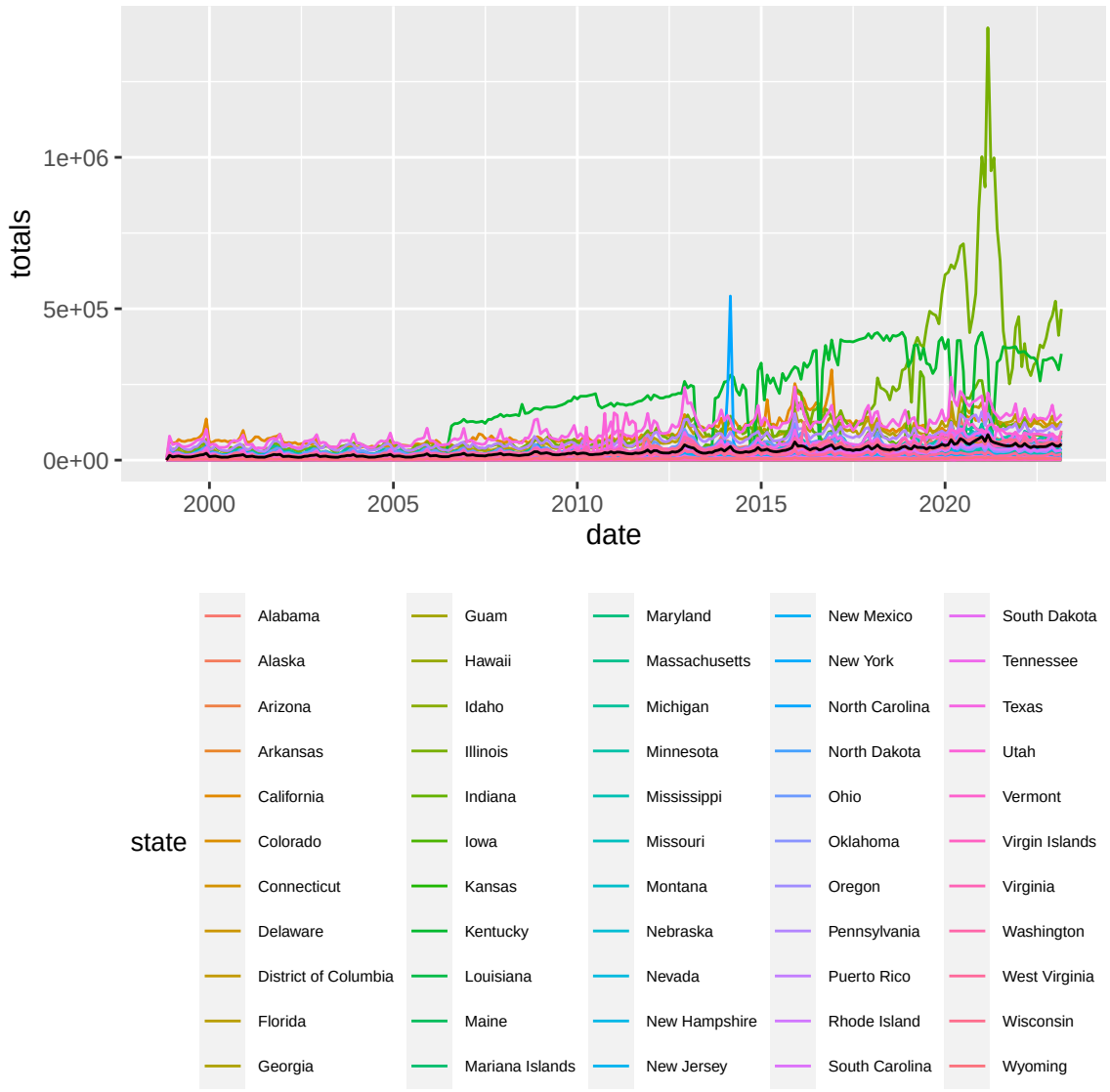
```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
dat <- firearm %>%
    mutate(date = ymd(paste(month, "01"))) %>%
    select(date, state, totals) %>%
    filter(complete.cases(.)) %>%
    arrange(date)

glimpse(dat)
```

```
## Rows: 16,115
## Columns: 3
## $ date   <date> 1998-11-01, 1998-11-01, 1998-11-01, 1998-11-01, 1998-11-01, 19~
## $ state  <chr> "Alabama", "Alaska", "Arizona", "Arkansas", "California", "Colo~
## $ totals <int> 1062, 145, 379, 589, 2101, 622, 80, 55, 0, 812, 62, 0, 28, 176,~
```

Plot of total background checks in each state as well as overall mean.

```
ggplot(dat) +
    geom_line(mapping=aes(x=date, y=totals, color=state)) +
    geom_line(dat %>% group_by(date) %>% summarise(totals=mean(totals)),
              mapping=aes(x=date, y=totals)) +
    theme(legend.position="bottom",
          legend.title = element_text(size=10),
          legend.text=element_text(size=6))
```



Plot of total background checks in each state as well as overall median on the *log* scale

```
ggplot(dat) +
    geom_line(mapping=aes(x=date, y=log(totals), color=state)) +
    geom_line(dat %>% group_by(date) %>% summarise(totals=median(totals)),
              mapping=aes(x=date, y=log(totals))) +
    theme(legend.position="bottom",
          legend.title = element_text( size=10),
          legend.text=element_text(size=6))
```

Time series decomposition of total background firearm checks

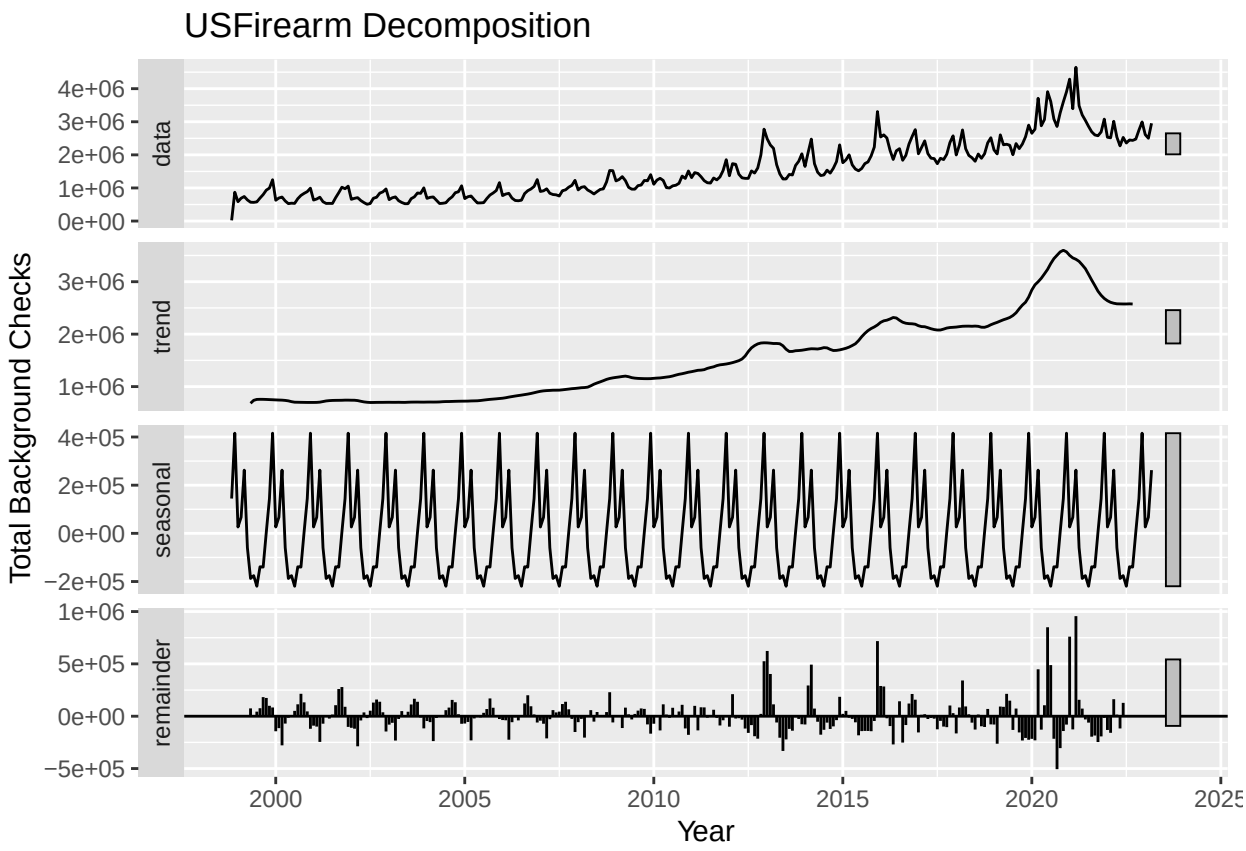
we are only interested in trend which is obtain after removing seasonal effect

```
dat <- dat |>
    group_by(date) |>
    summarise(totals = sum(totals))

glimpse(dat) # totals checks each month in the US
```

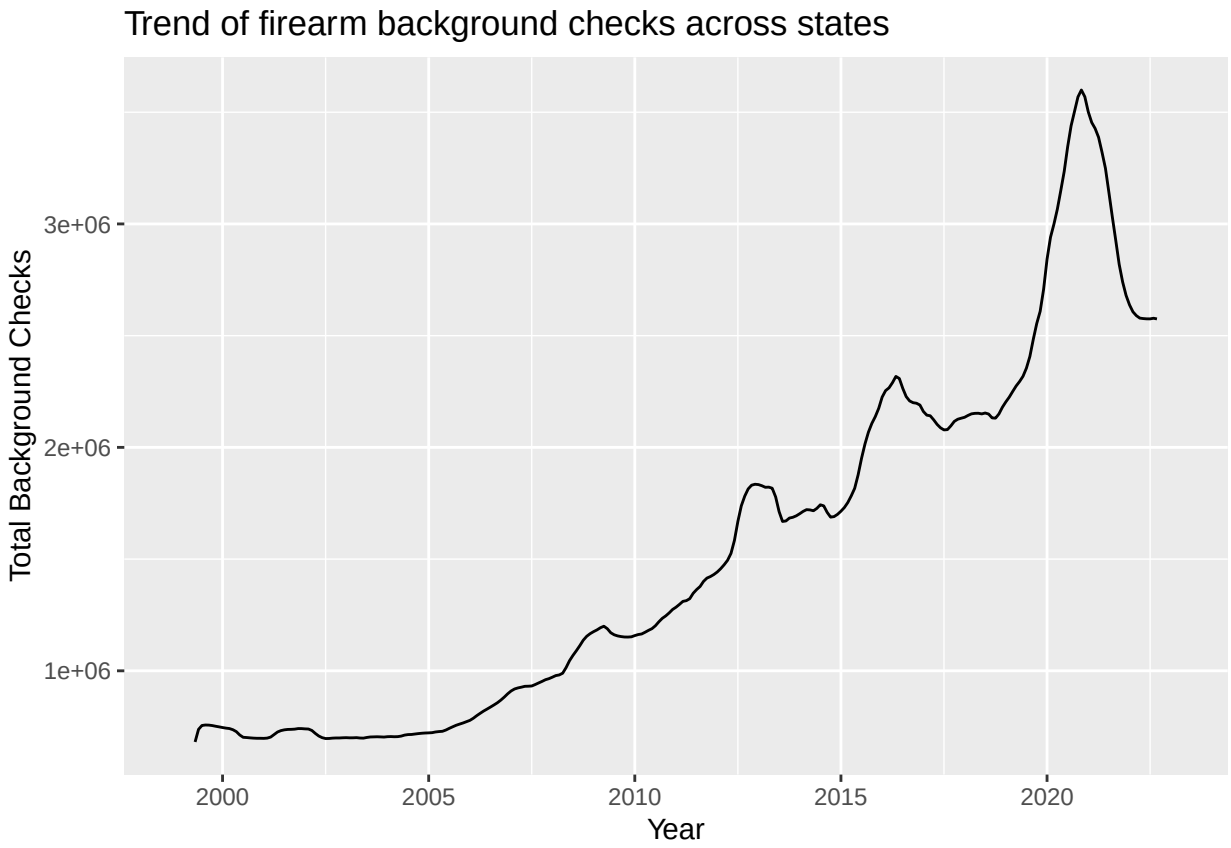
```
## Rows: 293
## Columns: 2
## $ date   <date> 1998-11-01, 1998-12-01, 1999-01-01, 1999-02-01, 1999-03-01, 19~
## $ totals <int> 21176, 870722, 585974, 690215, 741687, 638666, 569220, 564912, ~
```

```
firearm_ts <- ts(dat$totals,
             start = c(min(year(dat$date)), 11),
             end = c(max(year(dat$date)), 3),
             frequency = 12)
```

```
# Decomposing the model to obtain trend
fit_dcmp <- decompose(firearm_ts)

autoplot(fit_dcmp) +
  xlab("Year") +
  ylab("Total Background Checks") +
  ggtitle("USFirearm Decomposition")
```




```
autoplot(fit_dcmp$trend) +
    xlab("Year") +
    ylab("Total Background Checks") +
    ggtitle("Trend of firearm background checks across states")
```



The original data has strong seasonality (no cycles however).

After decomposing the time series, we have obtain the trend across the states.

Next I fit an ARIMA model that I can use to forecast ahead.

```
# fitting an ARIMA models to our time series
fit_arima <- auto.arima(firearm_ts)
summary(fit_arima)
```

```
## Series: firearm_ts
## ARIMA(2,0,1)(0,1,1)[12] with drift
##
## Coefficients:
##          ar1     ar2     ma1     sma1     drift
##       0.5792  0.2797  0.1257  -0.7057  8396.956
## s.e.  0.6333  0.5348  0.6641   0.0472  2319.809
##
## sigma^2 = 3.419e+10:  log likelihood = -3808.66
## AIC=7629.31   AICc=7629.62   BIC=7651.14
##
## Training set error measures:
##                      ME   RMSE    MAE       MPE     MAPE      MASE        ACF1
## Training set -2852.467 179455 106342 -1.099412 6.127842 0.4546183 -0.01925913
```

```
# forecasting for the next 24 months
forecast_arima <- forecast(fit_arima, h = 24)

autoplot(forecast_arima) +
```

```
  xlab("Year") +
  ylab("Total Background Checks") +
  ggtitle("24 Months Firearm Background Checks Forecast (ARIMA) ")
```