Homework1

xx,Roman Krass

2024-24-03

Part 1

```
##         DATE UNEMPLOY       DATE    POP       DATE PSAVERT
## 1 1948-01-01     2034 1952-01-01 156309 1959-01-01    11.3
## 2 1948-02-01     2328 1952-02-01 156527 1959-02-01    10.6
## 3 1948-03-01     2399 1952-03-01 156731 1959-03-01    10.3
```

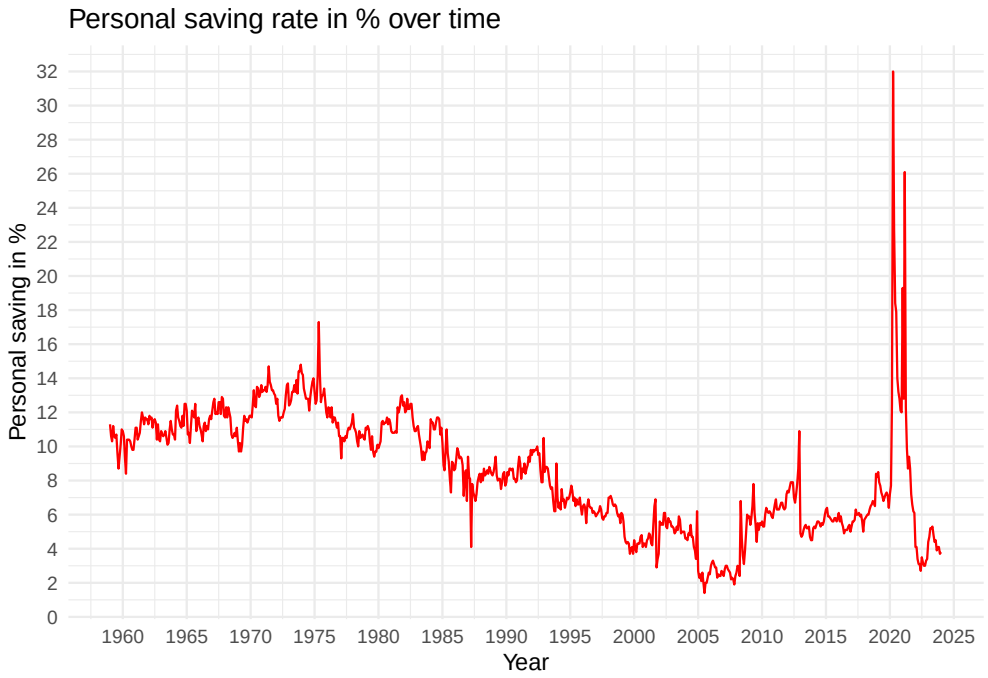
When loooking into the different datasets, we can see that the data doesn't have the same time period. Because of that when combining the three datasets there would be some NAs if we just would join the data. We can solve this by using the smallest common date range of the three datasets.

```
# Combine the three datasets to one tsibble with the DATE starting from 1959-01-01
data <- unemploy %>%
    filter(DATE >= "1959-01-01") %>%
  left_join(population, by = "DATE") %>%
  left_join(personalSavingRate, by = "DATE") %>%
  rename(unemployment = UNEMPLOY, population = POP, savingRate = PSAVERT) %>%
  as_tsibble(index = DATE)
```

```
# Plot the unemployment over time
ggplot(data, aes(x = DATE)) +
  geom_line(aes(y = unemployment), color = "red") +
    labs(title = "Unemployment rate over time",
        y = "Unemployment in thousands of people",
        x = "Year") +
    scale_x_date(date_breaks = "5 years", date_labels = "%Y") +
     theme_minimal()
```

In the plot we can see that the number of unemployed people has a lot of variation over time. But in this plot we just see the number of people that are unemployed. It would be interesting to see the unemployment rate in percent. What also can be seen very clearly is the COVID-19 pandemic in 2020. The number of unemployed people increased a lot in this year. In the plor we can also see an increase in the number of unemployed people in the years 2008 and 2009. This is the financial crisis that happened in these years.

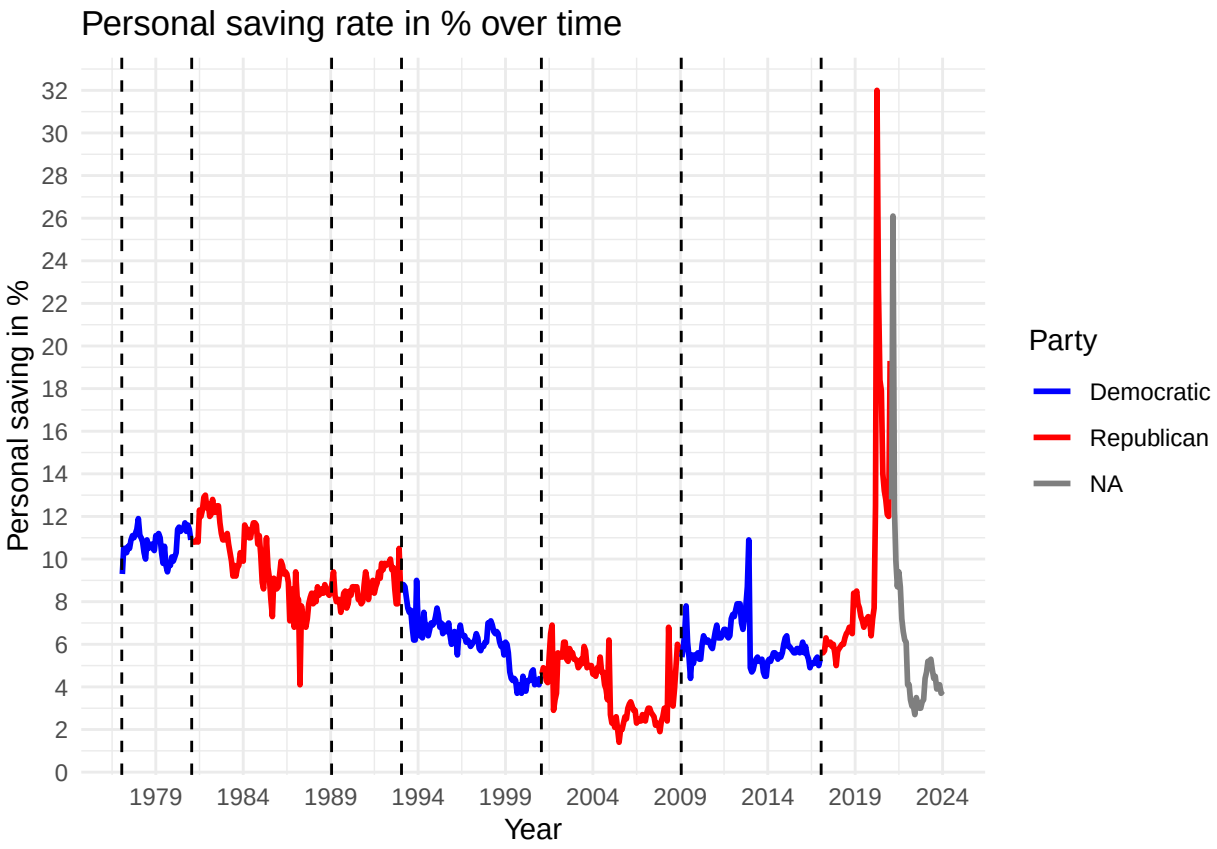
```
# Plot the personal saving rate over time
ggplot(data, aes(x = DATE)) +
  geom_line(aes(y = savingRate), color = "red") +
    labs(title = "Personal saving rate in % over time",
         y = "Personal saving in %",
         x = "Year") +
    scale_x_date(date_breaks = "5 years", date_labels = "%Y") +
    scale_y_continuous(breaks = seq(0, , 35, by = 2)) +
    theme_minimal()
```

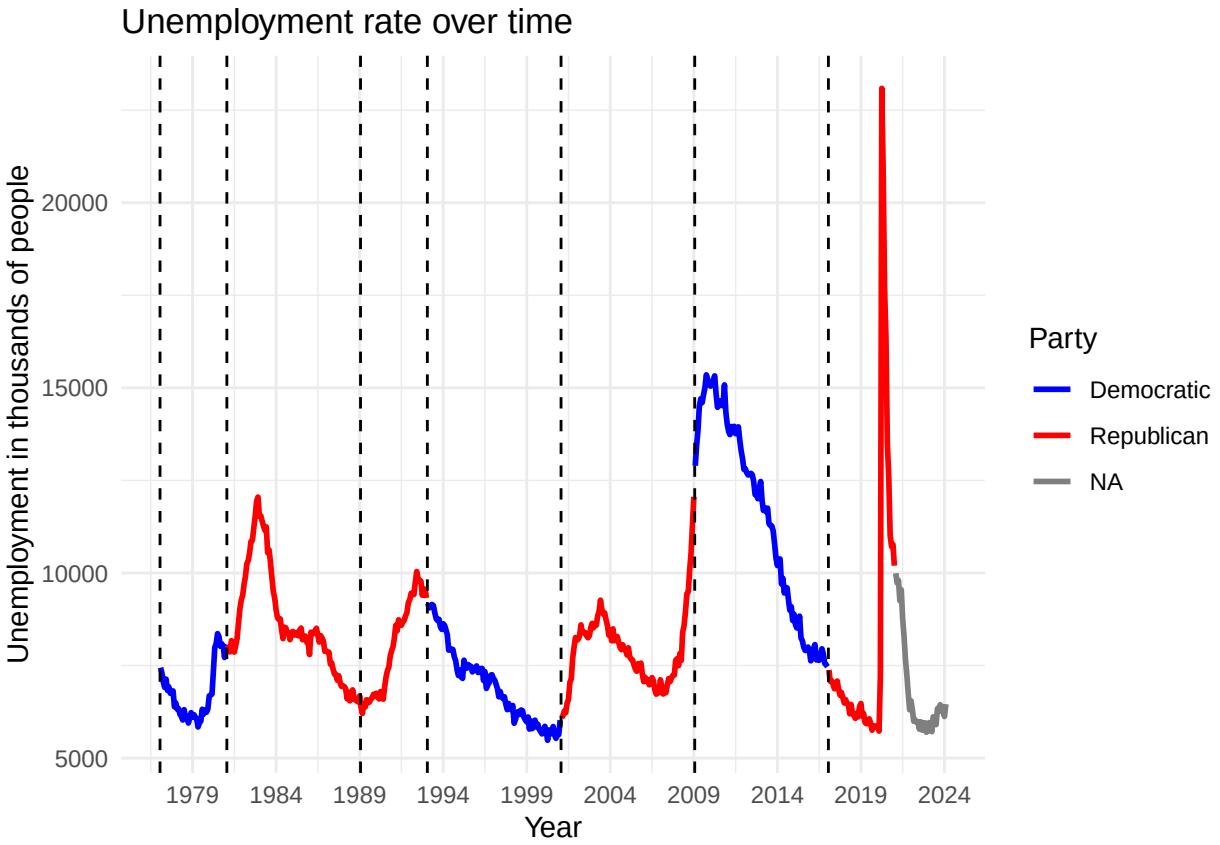


In the plot we can see that the personal saving rate decreased from about 1976 to 2007 from about 13% to 2%. After that the saving rate increased again to about 7% in 2020. Something extreme in the plot is the increase of the saving rate during the COVID-19 pandemic from 2019 to 2022. The saving rate increased from about 7% to 32% in 2020. This is a huge increase in the saving rate. After that the saving rate had a big de and increase. After the pandemic the saving rate dropped to about 3%. This makes sense because the people saved a lot of money during the pandemic and after the pandemic they started to spend the money again.

Part 2

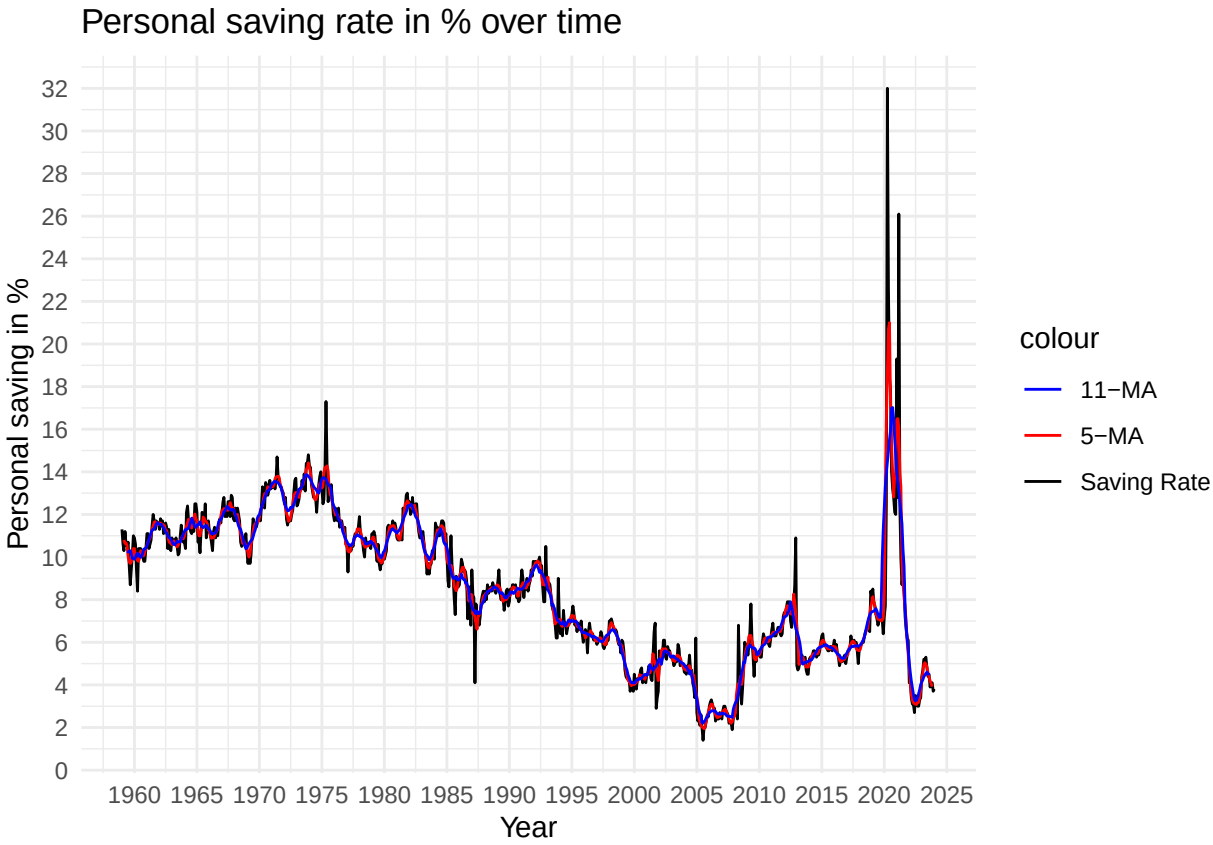
We would like to visualize how the ruling party and the election year affect unemployment and personal saving rate. Recreate the timeplots with the ruling parties shown by their color. Indicate also the election years on your plots. Start from Carter's presidency.


Personal saving rate in % over time
Personal saving in %
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
1979
1984
1989
1994
1999
2004
2009
2014
2019
2024
Year
Party
Democratic
Republican
NA

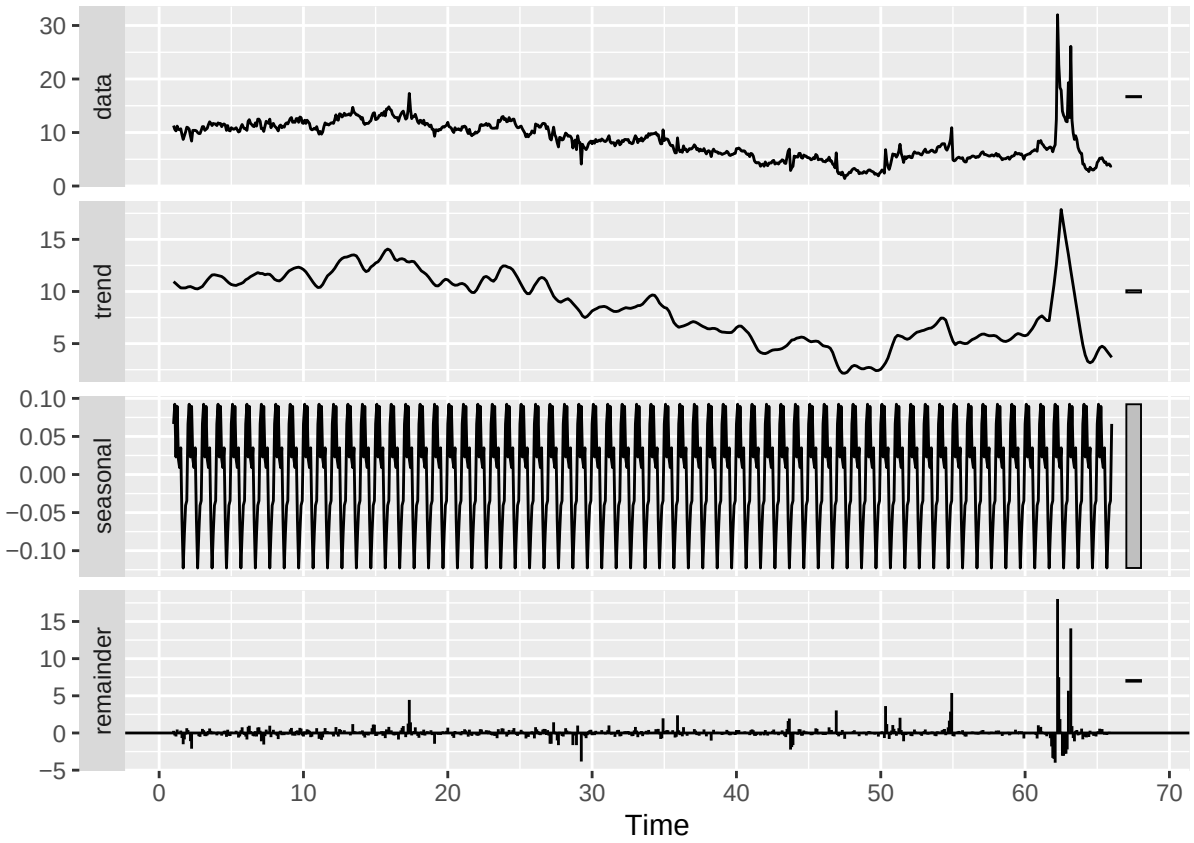


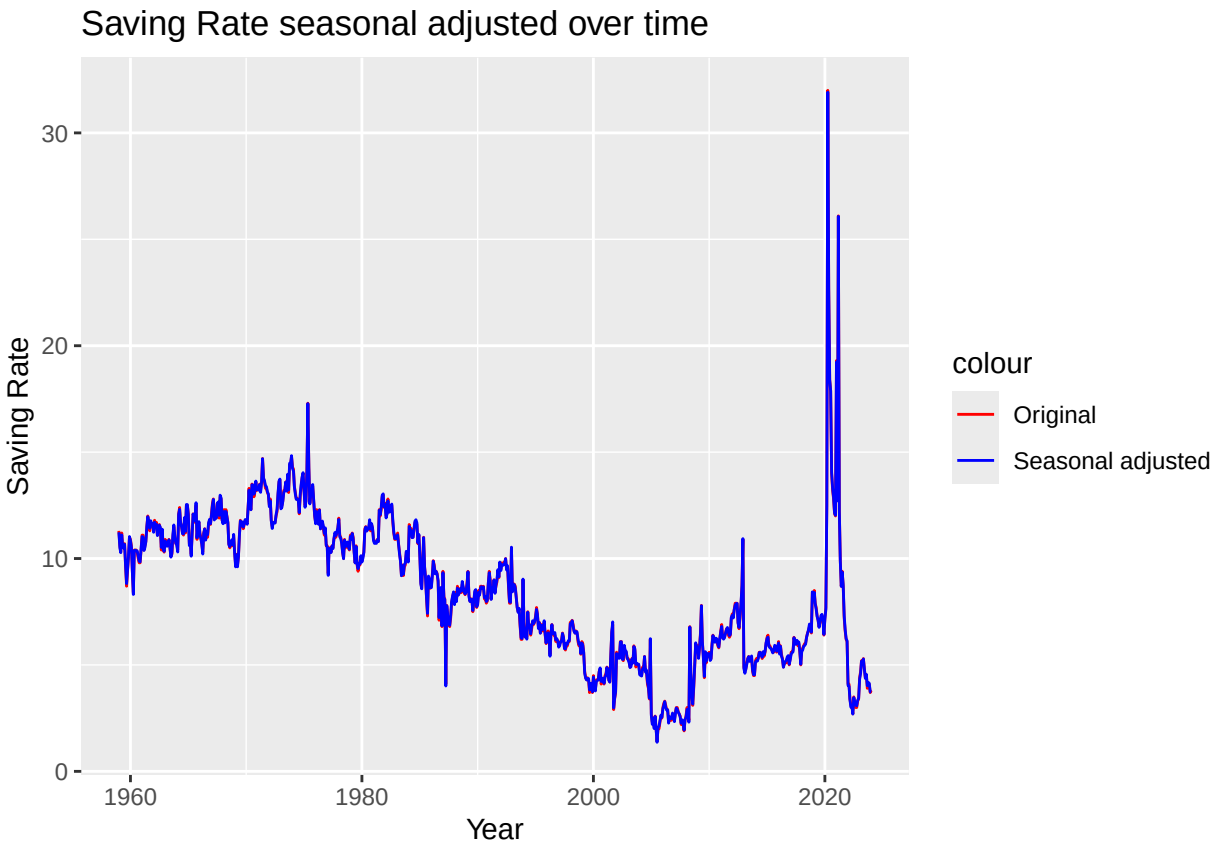
Part 3

Decompose the personal saving rate series using a centered moving average. Graph the season-adjusted data. Explain whether the remainder is a white noise. If you believe that a transformation could be useful before decomposing, then please do it. Compare your decomposition to the STL decomposition.


Personal saving rate in % over time
Personal saving in %
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025
Year
colour
11−MA
5−MA
Saving Rate

```
## [1] 1
```


data
trend
seasonal
remainder
Time


Saving Rate seasonal adjusted over time
Saving Rate
30
20
10
0
1960
1980
2000
2020
Year
colour
Original
Seasonal adjusted