Homework1

xx,Roman Krass

2024-24-03

```
## # A tibble: 3 x 2
##   DATE       UNEMPLOY
##   <date>        <dbl>
## 1 1948-01-01     2034
## 2 1948-02-01     2328
## 3 1948-03-01     2399
```

```
## # A tibble: 3 x 2
##   DATE          POP
##   <date>      <dbl>
## 1 1952-01-01 156309
## 2 1952-02-01 156527
## 3 1952-03-01 156731
```

```
## # A tibble: 3 x 2
##   DATE       PSAVERT
##   <date>       <dbl>
## 1 1959-01-01    11.3
## 2 1959-02-01    10.6
## 3 1959-03-01    10.3
```

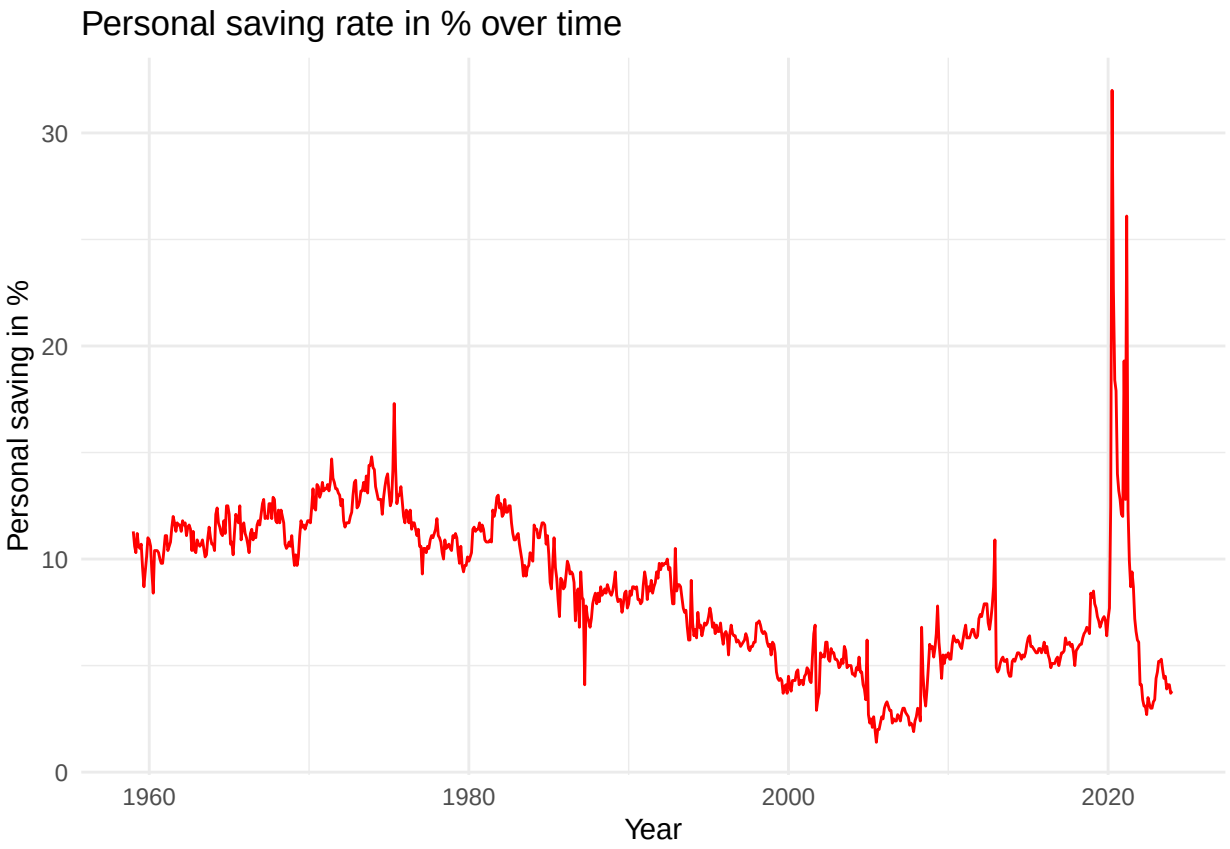
When loooking into the different datasets, we can see that the data doesn't have the same time period. Because of that when combining the three datasets there would be some NAs if we just would join the data. We can solve this by using the smallest common date range of the three datasets.

```
# 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") +
  theme_minimal() +
    labs(title = "Unemployment rate over time",
        y = "Unemployment in ",
        x = "Year")
```


```
# Plot the personal saving rate over time
ggplot(data, aes(x = DATE)) +
  geom_line(aes(y = savingRate), color = "red") +
  theme_minimal() +
    labs(title = "Personal saving rate in % over time",
         y = "Personal saving in %",
         x = "Year")
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_line()`).
```


Personal saving rate in % over time
Personal saving in %
30
20
10
0
1960
1980
2000
2020
Year