Week 3 Assignment

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.5.1     ✔ tibble    3.2.1
✔ lubridate 1.9.3     ✔ tidyr     1.3.1
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to
become errors
restatapi: - version 0.23.1
           - config file with the API version 2 loaded from GitHub (the 'current' API
version number is 2).
           - 2 from the 8 cores are used for parallel computing, can be changed with
'options(restatapi_cores=...)'
           - 'auto' method will be used for file download, can be changed with
'options(restatapi_dmethod=...)'
           - the Table of contents (TOC) was not pre-loaded into the deafult cache
('.restatapi_env').
```

Assignment 3

In this assignment, we're adding a few new skills and working on building more attractive and clearer data visualizations that show relationshiops between variables. Along the way, we're going to continue to learn about diversity within the European Union.

Joining Data

Often, we're going to want to join data from different sources. In this case, I've created a table that has some basic information about European Union Member States as well as some countries that have connections to the EU and therefore are included the Eurostat data. I've provided the files "ESNames.csv".

```
        ESNames <-read_csv('ESNames.csv', col_types = "ffffnf") # You may need to put a
full path to the file on your computer. The strange code at the end is to specify the
column types. (factor, factor, factor, factor, numeric, factor)

        summary(ESNames)
```

```
       CODE            countryname         Membership  EA20        AccYr
 EU27_2020: 1   European Union: 1   Group        : 3   No :37   Min.   :1957
 BE       : 1   Belgium       : 1   EUMember_2020:27   Yes:20   1st Qu.:1973
 BG       : 1   Bulgaria      : 1   NonMember    :17            Median :1995
 CZ       : 1   Czechia       : 1   EFTA         : 4            Mean   :1989
```

```
DK       : 1   Denmark       : 1   EU_Candidate : 6           3rd Qu.:2004
DE       : 1   Germany       : 1                              Max.   :2013
(Other)  :51   (Other)       :51                              NA's   :30
   ColdWar
NATO   :17
WP     :11
Neutral: 8
Yugo   : 6
NA's   :15
```

This dataset has a few variables that we can use to join with the Eurostat data.

CODE is the Eurostat code for the country or group of countries **countryname** is a more readable name for the country or group of countries **Membership** is the type of membership the country has relative to the EU **EA20** is a binary variable indicating whether the country is in the Euro Area **AccYr** is the year the country joined the EU (if it is a member) **ColdWar** identifies whether the country was part of NATO, the Warsaw Pact, Neutral, or part of Yugoslavia during the Cold War

```
GDP <- get_eurostat_data('tipsna40')
SB <- get_eurostat_data('tps00107')
```

We read in two datasets from Eurostat, one with GDP information (GDP) and one with information on spending for Social Benefits (SB).

```
summary(GDP)
```

```
           unit     na_item         geo           time         values
CLV15_EUR_HAB:835   B1GQ:835   AT     : 29   2000   : 29   Min.   :  3190
                               BE     : 29   2001   : 29   1st Qu.: 12850
                               BG     : 29   2002   : 29   Median : 22500
                               CY     : 29   2003   : 29   Mean   : 26034
                               CZ     : 29   2004   : 29   3rd Qu.: 35315
                               DE     : 29   2005   : 29   Max.   :101170
                               (Other):661   (Other):661
```

```
summary(SB)
```

```
     unit          spdeps          geo           time          values
PPS_HAB:4032   DISA   : 448   AT     : 108   2016   : 351   Min.   :    0.0
               EXCLU  : 448   BE     : 108   2017   : 351   1st Qu.:  164.1
               FAM    : 448   BG     : 108   2018   : 351   Median :  474.7
               HOUSE  : 448   CH     : 108   2015   : 342   Mean   : 1445.7
               OLD    : 448   CY     : 108   2013   : 333   3rd Qu.: 1742.1
               SICK   : 448   CZ     : 108   2014   : 333   Max.   :15903.6
               (Other):1344   (Other):3384   (Other):1971
```

For GDP, the values represent GDP per Capita in Euros for each country and year. For SB, the values represent social benefits per capita for each country, year and type of benefit (OLD, SICK, FAM, UNEMP, DISAB, HOUS, and OTH).

Homework

For each question, create a figure that explores the question and write a brief explanation of what you see. Clearly label your plots. Submit the .qmd file and the .html file to Canvas.

1. Do countries who joined the EU earlier tend to be wealthier or poorer than those who joined later? Create a plot that shows the relationship.

2. Do former communist countries spend more or less on social benefits for the old and sick than other countries? Create a plot that shows the relationship.(spdeps in SB has OLD and SICK)

3. Do wealthier countries in the EU spend more on social benefits for families (FAM) than poorer countries? Create a plot that shows the relationship.