Lab 1

```r
library(tidyverse)
```

```
-- Attaching packages -------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.4.0      v purrr   0.3.5
v tibble  3.1.8      v dplyr   1.0.10
v tidyr   1.2.1      v stringr 1.5.0
v readr   2.1.3      v forcats 0.5.2
```

```
Warning: package 'ggplot2' was built under R version 4.2.2
```

```
Warning: package 'readr' was built under R version 4.2.2
```

```
Warning: package 'purrr' was built under R version 4.2.2
```

```
Warning: package 'stringr' was built under R version 4.2.2
```

```
-- Conflicts ----------------------------------------- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

```r
dm <- read_table("https://www.prdh.umontreal.ca/BDLC/data/ont/Mx_1x1.txt", skip = 2, col_t
```

```
Warning: 494 parsing failures.
row    col                   expected actual
108 Female no trailing characters      . 'https://www.prdh.umontreal.ca/BDLC/data/ont/Mx_1x1
109 Female no trailing characters      . 'https://www.prdh.umontreal.ca/BDLC/data/ont/Mx_1x1
110 Female no trailing characters      . 'https://www.prdh.umontreal.ca/BDLC/data/ont/Mx_1x1
110 Male   no trailing characters      . 'https://www.prdh.umontreal.ca/BDLC/data/ont/Mx_1x1
```

```
110 Total  no trailing characters      . 'https://www.prdh.umontreal.ca/BDLC/data/ont/Mx_1x1
... ...... ..................... ...... ...............................................
See problems(...) for more details.
```

```
head(dm)
```

```
# A tibble: 6 x 5
   Year Age    Female    Male    Total
  <dbl> <chr>  <dbl>    <dbl>    <dbl>
1  1921 0     0.0978  0.129    0.114
2  1921 1     0.0129  0.0144   0.0137
3  1921 2     0.00521 0.00737 0.00631
4  1921 3     0.00471 0.00457 0.00464
5  1921 4     0.00461 0.00433 0.00447
6  1921 5     0.00372 0.00361 0.00367
```

Lab Exercises

1. Plot the ratio of male to female mortality rates over time for ages 10,20,30 and 40 (different color for each age) and change the theme
2. Find the age that has the highest female mortality rate each year
3. Use the `summarize(across())` syntax to calculate the standard deviation of mortality rates by age for the Male, Female and Total populations.
4. The Canadian HMD also provides population sizes over time (https://www.prdh.umontreal.ca/BDLC/data Use these to calculate the population weighted average mortality rate separately for males and females, for every year. Make a nice line plot showing the result (with meaningful labels/titles) and briefly comment on what you see (1 sentence). Hint: `left_join` will probably be useful here.

Question 1

```
dm <- dm |>
  mutate(mf_ratio = Male/Female)
dp <- dm |>
  filter(Age==10|Age==20|Age==30|Age==40)
dp |>
  ggplot(aes(x = Year, y = mf_ratio, color = Age)) +
  geom_line()+theme_bw()+  labs(y = "the ratio of male to female mortality rates")
```
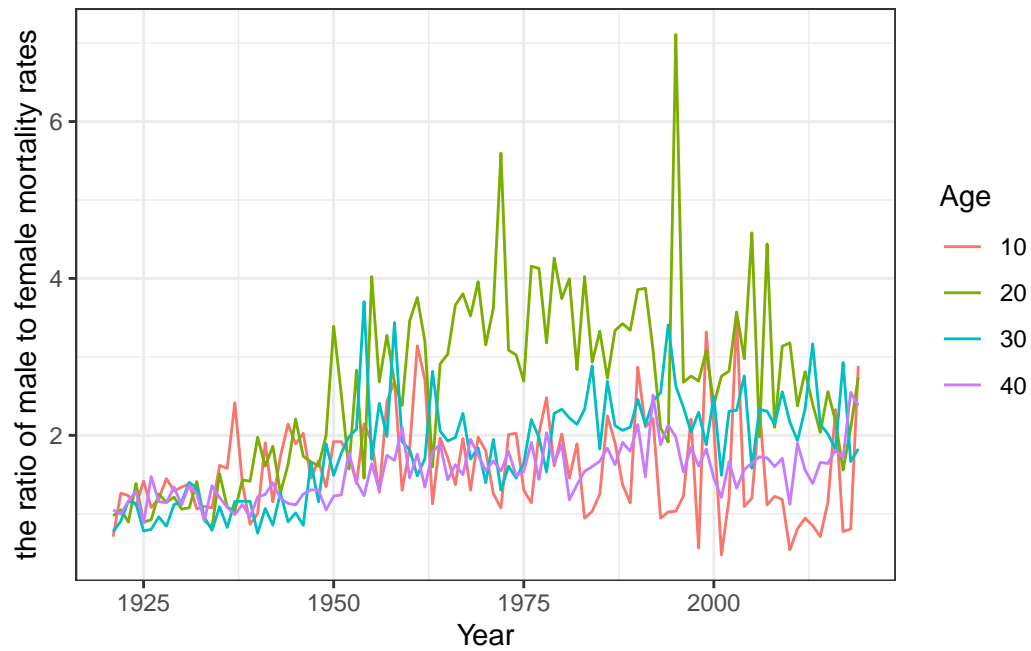
Question 2

```
dm |>
  group_by(Year)|>
  slice(which.max(Female))|>
  summarise(Age)
```

```
# A tibble: 99 x 2
    Year Age
   <dbl> <chr>
 1  1921 106
 2  1922 98
 3  1923 104
 4  1924 107
 5  1925 98
 6  1926 106
 7  1927 106
 8  1928 104
 9  1929 104
10  1930 105
# ... with 89 more rows
```

Question 3

```r
dm |>
  group_by(Age) |>
  summarize(across(c(Male,Female,Total),sd))
```

```
# A tibble: 111 x 4
   Age       Male     Female    Total
   <chr>     <dbl>     <dbl>     <dbl>
 1 0        0.0330   0.0256    0.0294
 2 1        0.00396  0.00352   0.00374
 3 10       0.000561 0.000474  0.000509
 4 100      0.138    0.0928    0.0729
 5 101      0.158    0.125     0.0995
 6 102      0.214    0.143     0.114
 7 103      0.371    0.252     0.208
 8 104      NA       0.449     0.363
 9 105      NA       NA        NA
10 106      NA       NA        NA
# ... with 101 more rows
```

Question 4

```r
ddd <- read_table("https://www.prdh.umontreal.ca/BDLC/data/ont/Population.txt", skip = 2,
```

```r
final <- left_join(dm,ddd,by = c("Year" = "Year", "Age" = "Age"))
final<-final|>
  group_by(Year)|>
  mutate(total_fe = sum(Female.y),total_ma = sum(Male.y),weight_fe = Female.y/total_fe,wei
  summarise(weighted.mean(Female.x,weight_fe,na.rm),weighted.mean(Male.x,weight_ma,na.rm))

colnames(final)[2] = "female"
colnames(final)[3] = "male"
```

```r
final |>
  ggplot(aes(x = Year))+ geom_line(aes(y = female ), color ="darkred")+geom_line(aes(y = m
```

The population weighted average mortality rate separately for
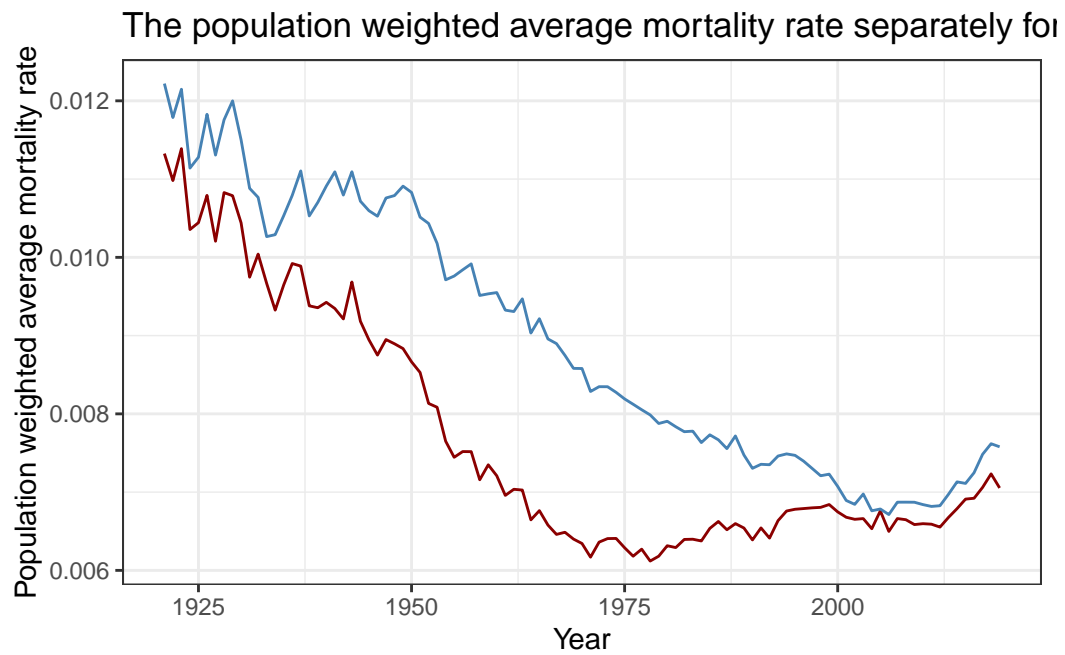
Comment: The male has a greater population weighted average mortality rate than females.