Course 1 Section 2.19 - Mutate

Jiaying Wu

14/09/2020

```
library(tidyverse)
library(here)
tb_long <- read_rds(here("data","tb_long.rds"))
tb_long
```

```
## # A tibble: 157,820 x 7
##    country     iso3   year type   sex   age_group count
##    <chr>       <chr> <dbl> <chr>  <chr> <chr>     <dbl>
##  1 Afghanistan AFG    1980 new_sp m     04           NA
##  2 Afghanistan AFG    1981 new_sp m     04           NA
##  3 Afghanistan AFG    1982 new_sp m     04           NA
##  4 Afghanistan AFG    1983 new_sp m     04           NA
##  5 Afghanistan AFG    1984 new_sp m     04           NA
##  6 Afghanistan AFG    1985 new_sp m     04           NA
##  7 Afghanistan AFG    1986 new_sp m     04           NA
##  8 Afghanistan AFG    1987 new_sp m     04           NA
##  9 Afghanistan AFG    1988 new_sp m     04           NA
## 10 Afghanistan AFG    1989 new_sp m     04           NA
## # ... with 157,810 more rows
```

Q1: Recode the sex variable to have values "male" and "female" instead of "m" and "f" (hint: use if_else)

```
tb_long <- tb_long %>% mutate(sex = if_else(sex == "m", "male", "female"))
tb_long
```

```
## # A tibble: 157,820 x 7
##    country     iso3   year type   sex   age_group count
##    <chr>       <chr> <dbl> <chr>  <chr> <chr>     <dbl>
##  1 Afghanistan AFG    1980 new_sp male  04           NA
##  2 Afghanistan AFG    1981 new_sp male  04           NA
##  3 Afghanistan AFG    1982 new_sp male  04           NA
##  4 Afghanistan AFG    1983 new_sp male  04           NA
##  5 Afghanistan AFG    1984 new_sp male  04           NA
##  6 Afghanistan AFG    1985 new_sp male  04           NA
##  7 Afghanistan AFG    1986 new_sp male  04           NA
##  8 Afghanistan AFG    1987 new_sp male  04           NA
##  9 Afghanistan AFG    1988 new_sp male  04           NA
## 10 Afghanistan AFG    1989 new_sp male  04           NA
## # ... with 157,810 more rows
```

Q2: Use spread to put the count values into the columns male and female, then create a new variable called diff which contains the difference between the male and female counts.

```
tb_wide <- tb_long %>%
  spread(key = "sex", value = "count") %>%
  mutate(diff = female - male)

tb_wide
```

```
## # A tibble: 78,910 x 8
##    country     iso3   year type   age_group female  male  diff
##    <chr>       <chr> <dbl> <chr>  <chr>       <dbl> <dbl> <dbl>
##  1 Afghanistan AFG    1980 new_sp 014            NA    NA    NA
##  2 Afghanistan AFG    1980 new_sp 04             NA    NA    NA
##  3 Afghanistan AFG    1980 new_sp 1524           NA    NA    NA
##  4 Afghanistan AFG    1980 new_sp 2534           NA    NA    NA
##  5 Afghanistan AFG    1980 new_sp 3544           NA    NA    NA
##  6 Afghanistan AFG    1980 new_sp 4554           NA    NA    NA
##  7 Afghanistan AFG    1980 new_sp 514            NA    NA    NA
##  8 Afghanistan AFG    1980 new_sp 5564           NA    NA    NA
##  9 Afghanistan AFG    1980 new_sp 65             NA    NA    NA
## 10 Afghanistan AFG    1980 new_sp u              NA    NA    NA
## # ... with 78,900 more rows
```