Lab-05A

Negar

2/17/2022

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1
```

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

```
lotr  <- read_csv("https://raw.githubusercontent.com/jennybc/lotr-tidy/master/data/lotr_tidy.csv") |>
  rename(Species = Race)
```

```
## Rows: 18 Columns: 4
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (3): Film, Race, Gender
## dbl (1): Words
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
lotr
```

```
## # A tibble: 18 x 4
##    Film                      Species Gender Words
##    <chr>                     <chr>   <chr>  <dbl>
##  1 The Fellowship Of The Ring Elf     Female  1229
##  2 The Fellowship Of The Ring Hobbit  Female    14
##  3 The Fellowship Of The Ring Man     Female     0
##  4 The Two Towers             Elf     Female   331
##  5 The Two Towers             Hobbit  Female     0
##  6 The Two Towers             Man     Female   401
##  7 The Return Of The King     Elf     Female   183
##  8 The Return Of The King     Hobbit  Female     2
##  9 The Return Of The King     Man     Female   268
## 10 The Fellowship Of The Ring Elf     Male     971
## 11 The Fellowship Of The Ring Hobbit  Male    3644
## 12 The Fellowship Of The Ring Man     Male    1995
## 13 The Two Towers             Elf     Male     513
## 14 The Two Towers             Hobbit  Male    2463
## 15 The Two Towers             Man     Male    3589
```

```
## 16 The Return Of The King     Elf     Male     510
## 17 The Return Of The King     Hobbit  Male    2673
## 18 The Return Of The King     Man     Male    2459
```

```
lotr_wide <- lotr |>
  pivot_wider(id_cols = c(Film, Gender),   # columns that we do not want to modify
              names_from = Species,
              values_from = Words)
lotr_wide
```

```
## # A tibble: 6 x 5
##   Film                       Gender   Elf Hobbit   Man
##   <chr>                      <chr>  <dbl>  <dbl> <dbl>
## 1 The Fellowship Of The Ring Female  1229     14     0
## 2 The Two Towers             Female   331      0   401
## 3 The Return Of The King     Female   183      2   268
## 4 The Fellowship Of The Ring Male     971   3644  1995
## 5 The Two Towers             Male     513   2463  3589
## 6 The Return Of The King     Male     510   2673  2459
```

```
lotr_wide |>
  pivot_longer(cols = Elf:Man,  # columns we want to pivot
               names_to = "Species",
               values_to = "Words"
               )
```

```
## # A tibble: 18 x 4
##    Film                       Gender Species Words
##    <chr>                      <chr>  <chr>   <dbl>
##  1 The Fellowship Of The Ring Female Elf      1229
##  2 The Fellowship Of The Ring Female Hobbit     14
##  3 The Fellowship Of The Ring Female Man         0
##  4 The Two Towers             Female Elf       331
##  5 The Two Towers             Female Hobbit      0
##  6 The Two Towers             Female Man       401
##  7 The Return Of The King     Female Elf       183
##  8 The Return Of The King     Female Hobbit      2
##  9 The Return Of The King     Female Man       268
## 10 The Fellowship Of The Ring Male   Elf       971
## 11 The Fellowship Of The Ring Male   Hobbit   3644
## 12 The Fellowship Of The Ring Male   Man      1995
## 13 The Two Towers             Male   Elf       513
## 14 The Two Towers             Male   Hobbit   2463
## 15 The Two Towers             Male   Man      3589
## 16 The Return Of The King     Male   Elf       510
## 17 The Return Of The King     Male   Hobbit   2673
## 18 The Return Of The King     Male   Man      2459
```

```
fam_dat <- tribble(
  ~family,  ~dob_child1,  ~dob_child2, ~gender_child1, ~gender_child2,
       1L, "1998-11-26", "2000-01-29",             1L,             2L,
       2L, "1996-06-22",           NA,             2L,             NA,
       3L, "2002-07-11", "2004-04-05",             2L,             2L,
       4L, "2004-10-10", "2009-08-27",             1L,             1L,
       5L, "2000-12-05", "2005-02-28",             2L,             1L,
)
```

```
fam_dat <- fam_dat |> mutate_at(vars(starts_with("dob")), parse_date)
fam_dat
```

```
## # A tibble: 5 x 5
##   family dob_child1 dob_child2 gender_child1 gender_child2
##    <int> <date>     <date>             <int>         <int>
## 1      1 1998-11-26 2000-01-29             1             2
## 2      2 1996-06-22 NA                     2            NA
## 3      3 2002-07-11 2004-04-05             2             2
## 4      4 2004-10-10 2009-08-27             1             1
## 5      5 2000-12-05 2005-02-28             2             1
```

```
fam_dat |>
  pivot_longer(
    cols = -family,
    names_to = c(".value", "child"),
    names_sep = "_",
    values_drop_na = TRUE
  )
```

```
## # A tibble: 9 x 4
##   family child  dob        gender
##    <int> <chr>  <date>      <int>
## 1      1 child1 1998-11-26      1
## 2      1 child2 2000-01-29      2
## 3      2 child1 1996-06-22      2
## 4      3 child1 2002-07-11      2
## 5      3 child2 2004-04-05      2
## 6      4 child1 2004-10-10      1
## 7      4 child2 2009-08-27      1
## 8      5 child1 2000-12-05      2
## 9      5 child2 2005-02-28      1
```

```
fam_dat_long <- fam_dat |>
  pivot_longer(
    cols = -family,
    names_to = c(".value", "child"),
    names_sep = "_",
    values_drop_na = TRUE
  ) |>
  mutate(child = stringr::str_replace(child, "child", "")) |>
  mutate(child = as.integer(child))
fam_dat_long
```

```
## # A tibble: 9 x 4
##   family child dob        gender
##    <int> <int> <date>      <int>
## 1      1     1 1998-11-26      1
## 2      1     2 2000-01-29      2
## 3      2     1 1996-06-22      2
## 4      3     1 2002-07-11      2
## 5      3     2 2004-04-05      2
## 6      4     1 2004-10-10      1
## 7      4     2 2009-08-27      1
## 8      5     1 2000-12-05      2
```

```
## 9      5     2 2005-02-28      1
```

```
fam_dat_long |>
  pivot_wider(id_cols     = family,
              names_from  = c(child, gender),
              names_prefix = "child",
              names_sep   = "_gender",
              values_from = dob)
```

```
## # A tibble: 5 x 5
##   family child1_gender1 child2_gender2 child1_gender2 child2_gender1
##    <int> <date>         <date>         <date>         <date>
## 1      1 1998-11-26     2000-01-29     NA             NA
## 2      2 NA             NA             1996-06-22     NA
## 3      3 NA             2004-04-05     2002-07-11     NA
## 4      4 2004-10-10     NA             NA             2009-08-27
## 5      5 NA             NA             2000-12-05     2005-02-28
```

```
fam_dat_long |>
  pivot_wider(id_cols     = family,
              names_from  = child,
              names_prefix = "child",
              names_sep   = "_",
              values_from = c(dob, gender))
```

```
## # A tibble: 5 x 5
##   family dob_child1 dob_child2 gender_child1 gender_child2
##    <int> <date>     <date>             <int>         <int>
## 1      1 1998-11-26 2000-01-29             1             2
## 2      2 1996-06-22 NA                     2            NA
## 3      3 2002-07-11 2004-04-05             2             2
## 4      4 2004-10-10 2009-08-27             1             1
## 5      5 2000-12-05 2005-02-28             2             1
```

```
fam_dat_long |>
  pivot_wider(id_cols     = family,
              names_from  = child,
              names_glue = "child{child}_{.value}",
              values_from = c(dob, gender))
```

```
## # A tibble: 5 x 5
##   family child1_dob child2_dob child1_gender child2_gender
##    <int> <date>     <date>             <int>         <int>
## 1      1 1998-11-26 2000-01-29             1             2
## 2      2 1996-06-22 NA                     2            NA
## 3      3 2002-07-11 2004-04-05             2             2
## 4      4 2004-10-10 2009-08-27             1             1
## 5      5 2000-12-05 2005-02-28             2             1
```