homework_03

07 On Your Own #2-3

2. Write a function to give choices about year, county, and variables

```r
acs_function <- function(year, variables, county){
    tidycensus::get_acs(
    year = year,
    state = "MN",
    geography = "tract",
    variables = variables,
    output = "wide",
    geometry = TRUE,
    county = county,
    show_call = TRUE
    )
}


acs_function(2021, c("B01003_001E", "B19013_001E"), "Hennepin")
```

```
Getting data from the 2017-2021 5-year ACS
```

```
Warning: * You have not set a Census API key. Users without a key are limited to 500
queries per day and may experience performance limitations.
i For best results, get a Census API key at
http://api.census.gov/data/key_signup.html and then supply the key to the
`census_api_key()` function to use it throughout your tidycensus session.
This warning is displayed once per session.
```

```
Downloading feature geometry from the Census website.  To cache shapefiles for use in future
```

```
Census API call: https://api.census.gov/data/2021/acs/acs5?get=B01003_001E%2CB01003_001M%2CB

  |
  |                                                                      |   0%
  |
  |=                                                                     |   2%
  |
  |==                                                                    |   3%
  |
  |===                                                                   |   5%
  |
  |====                                                                  |   6%
  |
  |======                                                                |   8%
  |
  |=======                                                               |  10%
  |
  |========                                                              |  11%
  |
  |=========                                                             |  13%
  |
  |==========                                                            |  14%
  |
  |===========                                                           |  16%
  |
  |============                                                          |  18%
  |
  |===============                                                       |  21%
  |
  |================                                                      |  22%
  |
  |==================                                                    |  25%
  |
  |===================                                                   |  27%
  |
  |====================                                                  |  28%
  |
  |======================                                                |  32%
  |
  |=============================                                         |  41%
  |
|==============================                                        |  43%
|
|===============================                                       |  45%
|
|================================                                      |  46%
|
|=====================================                                 |  52%
|
|======================================                                |  54%
|
|=======================================                               |  56%
|
|========================================                              |  57%
|
|=========================================                             |  59%
|
|==========================================                            |  60%
|
|===========================================                           |  62%
|
|=============================================                         |  64%
|
|==============================================                        |  65%
|
|=================================================                     |  70%
|
|===================================================                   |  73%
|
|====================================================                  |  75%
|
|=====================================================                 |  76%
|
|=======================================================               |  78%
|
|========================================================              |  80%
|
|=========================================================             |  81%
|
|==========================================================            |  83%
|
|============================================================          |  86%
|
|==============================================================        |  89%
```

```
  |
  |================================================================      |  91%
  |
  |=================================================================     |  92%
  |
  |===================================================================   |  96%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%
```

```
Simple feature collection with 329 features and 6 fields
Geometry type: MULTIPOLYGON
Dimension:     XY
Bounding box:  xmin: -93.76838 ymin: 44.78538 xmax: -93.17722 ymax: 45.24662
Geodetic CRS:  NAD83
First 10 features:
         GEOID                                          NAME B01003_001E
1  27053024300    Census Tract 243, Hennepin County, Minnesota        4744
2  27053110500   Census Tract 1105, Hennepin County, Minnesota        4969
3  27053024006 Census Tract 240.06, Hennepin County, Minnesota        2205
4  27053022801 Census Tract 228.01, Hennepin County, Minnesota        2481
5  27053026908 Census Tract 269.08, Hennepin County, Minnesota        6139
6  27053025401 Census Tract 254.01, Hennepin County, Minnesota        4428
7  27053108600   Census Tract 1086, Hennepin County, Minnesota        2947
8  27053026824 Census Tract 268.24, Hennepin County, Minnesota        4551
9  27053106000   Census Tract 1060, Hennepin County, Minnesota        3375
10 27053000102   Census Tract 1.02, Hennepin County, Minnesota        4896
   B01003_001M B19013_001E B19013_001M                       geometry
1          481       72240        5745 MULTIPOLYGON (((-93.31881 4...
2          651       80157        5307 MULTIPOLYGON (((-93.22237 4...
3          270      143125       22624 MULTIPOLYGON (((-93.35044 4...
4          359      133958       34619 MULTIPOLYGON (((-93.34793 4...
5          792      110246        3614 MULTIPOLYGON (((-93.39145 4...
6          648       68711       11097 MULTIPOLYGON (((-93.28347 4...
7          587       57470       15799 MULTIPOLYGON (((-93.24995 4...
8          483      127819       26964 MULTIPOLYGON (((-93.36073 4...
9          622       23492        5316 MULTIPOLYGON (((-93.25966 4...
10         597       59750       11634 MULTIPOLYGON (((-93.29919 4...
```

3. Use your function from (2) along with `map` and `list_rbind` to build a data set for Rice county for the years 2019-2021

```r
acs_rice <- 2019:2021 |>
  map(acs_function, variables = c("B01003_001E", "B19013_001E"), county = "Rice")
```

```
Getting data from the 2015-2019 5-year ACS

Downloading feature geometry from the Census website.  To cache shapefiles for use in future

Census API call: https://api.census.gov/data/2019/acs/acs5?get=B01003_001E%2CB01003_001M%2CB


  |
  |                                                                      |   0%
  |
  |===========                                                           |  16%
  |
  |=============                                                         |  18%
  |
  |==============                                                        |  20%
  |
  |===============                                                       |  22%
  |
  |================                                                      |  23%
  |
  |==================                                                    |  25%
  |
  |===================                                                   |  27%
  |
  |====================                                                  |  29%
  |
  |======================                                                |  32%
  |
  |========================                                              |  34%
  |
  |=========================                                             |  36%
  |
  |==========================                                            |  38%
  |
  |============================                                          |  39%
  |
  |=============================                                         |  41%
  |
|==============================                                        |  43%
|
|===============================                                       |  45%
|
|=================================                                     |  47%
|
|===================================                                   |  50%
|
|====================================                                  |  52%
|
|======================================                                |  54%
|
|=======================================                               |  56%
|
|========================================                              |  57%
|
|=========================================                             |  59%
|
|===========================================                           |  61%
|
|============================================                          |  63%
|
|=============================================                         |  64%
|
|==============================================                        |  66%
|
|================================================                      |  68%
|
|=================================================                     |  70%
|
|==================================================                    |  72%
|
|===================================================                   |  73%
|
|======================================================                |  77%
|
|=======================================================               |  79%
|
|==========================================================            |  82%
|
|=============================================================         |  88%
|
|================================================================      |  91%
```

```
  |
  |=================================================================     |  93%
  |
  |==================================================================    |  95%
  |
  |====================================================================  |  97%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%

Getting data from the 2016-2020 5-year ACS

Downloading feature geometry from the Census website.  To cache shapefiles for use in future

Census API call: https://api.census.gov/data/2020/acs/acs5?get=B01003_001E%2CB01003_001M%2CB


  |
  |                                                                      |   0%
  |
  |=                                                                     |   2%
  |
  |==                                                                    |   3%
  |
  |===                                                                   |   5%
  |
  |=====                                                                 |   8%
  |
  |======                                                                |   9%
  |
  |=======                                                               |  11%
  |
  |========                                                              |  12%
  |
  |==========                                                            |  14%
  |
  |===========                                                           |  15%
  |
  |============                                                          |  17%
  |
|=============                                                         |  18%
|
|===============                                                       |  21%
|
|=================                                                     |  24%
|
|===================                                                   |  27%
|
|====================                                                  |  28%
|
|=====================                                                 |  30%
|
|======================                                                |  31%
|
|========================                                              |  34%
|
|=========================                                             |  36%
|
|==========================                                            |  38%
|
|===========================                                           |  39%
|
|=============================                                         |  42%
|
|===============================                                       |  44%
|
|================================                                      |  45%
|
|=================================                                     |  47%
|
|==================================                                    |  48%
|
|===================================                                   |  49%
|
|====================================                                  |  51%
|
|=====================================                                 |  53%
|
|======================================                                |  54%
|
|=======================================                               |  56%
|
|=========================================                             |  59%
```

```
  |
  |==========================================                            |  60%
  |
  |==================================================                    |  71%
  |
  |====================================================                  |  74%
  |
  |=======================================================               |  78%
  |
  |========================================================              |  80%
  |
  |=========================================================             |  81%
  |
  |==========================================================            |  83%
  |
  |============================================================          |  86%
  |
  |=============================================================         |  87%
  |
  |==============================================================        |  89%
  |
  |===============================================================       |  90%
  |
  |================================================================      |  92%
  |
  |=================================================================     |  93%
  |
  |==================================================================    |  95%
  |
  |===================================================================   |  96%
  |
  |====================================================================  |  98%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%
```

```
Getting data from the 2017-2021 5-year ACS

Downloading feature geometry from the Census website.  To cache shapefiles for use in future

Census API call: https://api.census.gov/data/2021/acs/acs5?get=B01003_001E%2CB01003_001M%2CB
```

```
list_rbind(acs_rice)
```

```
         GEOID                                        NAME B01003_001E
1  27131070504 Census Tract 705.04, Rice County, Minnesota        3933
2  27131070400    Census Tract 704, Rice County, Minnesota        4511
3  27131070300    Census Tract 703, Rice County, Minnesota        4551
4  27131070503 Census Tract 705.03, Rice County, Minnesota        3348
5  27131070601 Census Tract 706.01, Rice County, Minnesota        3526
6  27131070800    Census Tract 708, Rice County, Minnesota        8101
7  27131070901 Census Tract 709.01, Rice County, Minnesota        5509
8  27131070700    Census Tract 707, Rice County, Minnesota        7165
9  27131070100    Census Tract 701, Rice County, Minnesota        7333
10 27131070602 Census Tract 706.02, Rice County, Minnesota        5211
11 27131070200    Census Tract 702, Rice County, Minnesota        5463
12 27131070902 Census Tract 709.02, Rice County, Minnesota        3160
13 27131070501 Census Tract 705.01, Rice County, Minnesota        4374
14 27131070501 Census Tract 705.01, Rice County, Minnesota        4272
15 27131070504 Census Tract 705.04, Rice County, Minnesota        3941
16 27131070801 Census Tract 708.01, Rice County, Minnesota        4456
17 27131070200    Census Tract 702, Rice County, Minnesota        5508
18 27131070701 Census Tract 707.01, Rice County, Minnesota        3057
19 27131070400    Census Tract 704, Rice County, Minnesota        4686
20 27131070300    Census Tract 703, Rice County, Minnesota        4737
21 27131070601 Census Tract 706.01, Rice County, Minnesota        3669
22 27131070102 Census Tract 701.02, Rice County, Minnesota        3786
23 27131070802 Census Tract 708.02, Rice County, Minnesota        3873
24 27131070702 Census Tract 707.02, Rice County, Minnesota        3872
25 27131070901 Census Tract 709.01, Rice County, Minnesota        5681
26 27131070503 Census Tract 705.03, Rice County, Minnesota        3185
27 27131070902 Census Tract 709.02, Rice County, Minnesota        2992
28 27131070101 Census Tract 701.01, Rice County, Minnesota        3428
29 27131070602 Census Tract 706.02, Rice County, Minnesota        5406
30 27131070902 Census Tract 709.02, Rice County, Minnesota        3212
31 27131070601 Census Tract 706.01, Rice County, Minnesota        3775
32 27131070503 Census Tract 705.03, Rice County, Minnesota        3035
33 27131070702 Census Tract 707.02, Rice County, Minnesota        3738
34 27131070901 Census Tract 709.01, Rice County, Minnesota        5858
35 27131070801 Census Tract 708.01, Rice County, Minnesota        4618
36 27131070501 Census Tract 705.01, Rice County, Minnesota        4242
37 27131070300    Census Tract 703, Rice County, Minnesota        4657
38 27131070200    Census Tract 702, Rice County, Minnesota        5419
39 27131070400    Census Tract 704, Rice County, Minnesota        4380
```

```
40 27131070701 Census Tract 707.01, Rice County, Minnesota        3028
41 27131070504 Census Tract 705.04, Rice County, Minnesota        3917
42 27131070101 Census Tract 701.01, Rice County, Minnesota        3417
43 27131070802 Census Tract 708.02, Rice County, Minnesota        3944
44 27131070102 Census Tract 701.02, Rice County, Minnesota        4201
45 27131070602 Census Tract 706.02, Rice County, Minnesota        5354
   B01003_001M B19013_001E B19013_001M                       geometry
1          273       63989        9273 MULTIPOLYGON (((-93.19137 4...
2          168       85952        2758 MULTIPOLYGON (((-93.40564 4...
3          190       78343        4242 MULTIPOLYGON (((-93.52521 4...
4          245       92321       14200 MULTIPOLYGON (((-93.16075 4...
5          333       50368        9979 MULTIPOLYGON (((-93.17615 4...
6          465       48403        7679 MULTIPOLYGON (((-93.29819 4...
7          456       44417       10552 MULTIPOLYGON (((-93.30904 4...
8          414       67868        9422 MULTIPOLYGON (((-93.27265 4...
9          326       91667        8106 MULTIPOLYGON (((-93.52452 4...
10         310       64479       12376 MULTIPOLYGON (((-93.22644 4...
11         177      101359        4104 MULTIPOLYGON (((-93.5246 44...
12         410       45230       12887 MULTIPOLYGON (((-93.30888 4...
13         270       66188        9179 MULTIPOLYGON (((-93.16981 4...
14         316       64792       13256 MULTIPOLYGON (((-93.16981 4...
15         536       63500        7351 MULTIPOLYGON (((-93.1909 44...
16         703       67625       23325 MULTIPOLYGON (((-93.29829 4...
17         473      104011        5648 MULTIPOLYGON (((-93.5246 44...
18         218       73750       13139 MULTIPOLYGON (((-93.26704 4...
19         296       86094        3438 MULTIPOLYGON (((-93.40564 4...
20         244       79068        4902 MULTIPOLYGON (((-93.52518 4...
21         525       52936       10436 MULTIPOLYGON (((-93.17615 4...
22         199       96023       13649 MULTIPOLYGON (((-93.44292 4...
23         437       63924        8715 MULTIPOLYGON (((-93.28272 4...
24         425       49811       16864 MULTIPOLYGON (((-93.27265 4...
25         566       51595        9615 MULTIPOLYGON (((-93.30904 4...
26         341      100516       11630 MULTIPOLYGON (((-93.16075 4...
27         440       46750       15457 MULTIPOLYGON (((-93.30888 4...
28         295      100563       15809 MULTIPOLYGON (((-93.52452 4...
29         377       62078        5270 MULTIPOLYGON (((-93.22644 4...
30         421       47059       15456 MULTIPOLYGON (((-93.30888 4...
31         435       56319        4333 MULTIPOLYGON (((-93.17615 4...
32         321      105952        8429 MULTIPOLYGON (((-93.16075 4...
33         409       57126       13968 MULTIPOLYGON (((-93.27265 4...
34         714       47344        9579 MULTIPOLYGON (((-93.30904 4...
35         622       61193       23977 MULTIPOLYGON (((-93.29829 4...
36         380       79063       15272 MULTIPOLYGON (((-93.16981 4...
37          296       83911        7244 MULTIPOLYGON (((-93.52522 4...
38          520      111711       10313 MULTIPOLYGON (((-93.5246 44...
39          274       90179        4919 MULTIPOLYGON (((-93.40564 4...
40          358       82500       20934 MULTIPOLYGON (((-93.26775 4...
41          537       67219        9805 MULTIPOLYGON (((-93.1909 44...
42          270      108490        1768 MULTIPOLYGON (((-93.52452 4...
43          462       63679       12261 MULTIPOLYGON (((-93.28274 4...
44          199       85789       20094 MULTIPOLYGON (((-93.44292 4...
45          359       63835        4805 MULTIPOLYGON (((-93.22644 4...
```

OMDB example

```
#myapikey01 <- Sys.getenv("OMDB_KEY")
#url <- str_c("https://omdbapi.com/?t=Moonlight&y=2016&apikey=", myapikey01)
#I am having trouble getting this code to work using Sys.setenv and Sys.getenv through the


url <- "https://omdbapi.com/?t=Moonlight&y=2016&apikey=235e12da"

moonlight <- GET(url)
moonlight
```

```
Response [https://omdbapi.com/?t=Moonlight&y=2016&apikey=235e12da]
  Date: 2024-10-23 02:10
  Status: 200
  Content-Type: application/json; charset=utf-8
  Size: 965 B
```

```
details <- content(moonlight, "parse")
details
```

```
$Title
[1] "Moonlight"

$Year
[1] "2016"

$Rated
[1] "R"
```

```
$Released
[1] "18 Nov 2016"

$Runtime
[1] "111 min"

$Genre
[1] "Drama"

$Director
[1] "Barry Jenkins"

$Writer
[1] "Barry Jenkins, Tarell Alvin McCraney"

$Actors
[1] "Mahershala Ali, Naomie Harris, Trevante Rhodes"

$Plot
[1] "A young African-American man grapples with his identity and sexuality while experiencing

$Language
[1] "English"

$Country
[1] "United States"

$Awards
[1] "Won 3 Oscars. 233 wins & 310 nominations total"

$Poster
[1] "https://m.media-amazon.com/images/M/MV5BNzQxNTIyODAxMV5BMl5BanBnXkFtZTgwNzQyMDA3OTE@._V

$Ratings
$Ratings[[1]]
$Ratings[[1]]$Source
[1] "Internet Movie Database"

$Ratings[[1]]$Value
[1] "7.4/10"
```

```
$Ratings[[2]]
$Ratings[[2]]$Source
[1] "Rotten Tomatoes"

$Ratings[[2]]$Value
[1] "98%"


$Ratings[[3]]
$Ratings[[3]]$Source
[1] "Metacritic"

$Ratings[[3]]$Value
[1] "99/100"



$Metascore
[1] "99"

$imdbRating
[1] "7.4"

$imdbVotes
[1] "336,072"

$imdbID
[1] "tt4975722"

$Type
[1] "movie"

$DVD
[1] "N/A"

$BoxOffice
[1] "$27,854,932"

$Production
[1] "N/A"

$Website
[1] "N/A"
```

```
$Response
[1] "True"
```

```
movies <- c("Moonlight", "Little+Women", "Call+Me+By+Your+Name", "Fences", "In+The+Heights

omdb <- tibble(Title = character(), Awards = character(), Runtime = double(), BoxOffice =

# Use for loop to run through API request process 5 times,
#   each time filling the next row in the tibble
#  - can do max of 1000 GETs per day
for(i in 1:5) {
  #url <- str_c("http://www.omdbapi.com/?t=",movies[i],
   #            "&apikey=", myapikey01)
  url <- str_c("http://www.omdbapi.com/?t=",movies[i],
               "&apikey=235e12da")
  Sys.sleep(0.5)
  onemovie <- GET(url)
  details <- content(onemovie, "parse")
  omdb[i,1] <- details$Title
  omdb[i,2] <- details$Awards
  omdb[i,3] <- parse_number(details$Runtime)
  omdb[i,4] <- parse_number(details$BoxOffice)
  omdb[i,5] <- parse_number(details$Year)
}

omdb
```

```
# A tibble: 5 x 5
  Title                Awards                             Runtime BoxOffice  Year
  <chr>                <chr>                                <dbl>     <dbl> <dbl>
1 Moonlight            Won 3 Oscars. 233 wins & 310 nom~      111  27854932  2016
2 Little Women         Won 1 Oscar. 78 wins & 239 nomin~      135 108101214  2019
3 Call Me by Your Name Won 1 Oscar. 104 wins & 262 nomi~      132  18095701  2017
4 Fences               Won 1 Oscar. 60 wins & 122 nomin~      139  57682904  2016
5 In the Heights       11 wins & 57 nominations               143  29975167  2021
```

08 On Your Own #2.2-2.4

2) Organize your `rvest` code from (1) into functions from the `polite` package.

```r
session <- bow("https://www.hockey-reference.com/teams/MIN/2001.html", force = TRUE)
result <- scrape(session) |>
 html_nodes(css = "table") |>
  html_table(header = TRUE, fill = TRUE)
```

```
No encoding supplied: defaulting to UTF-8.
```

```r
player_tibble <- result[[4]]
player_tibble
```

```
# A tibble: 40 x 22
   ``    ``    ``    ``    ``    Scoring Scoring Scoring ``    ``    Goals Goals
   <chr> <chr> <chr> <chr> <chr> <chr>   <chr>   <chr>   <chr> <chr> <chr> <chr>
 1 Rk    Play~ Age   Pos   GP    G       A       PTS     +/-   PIM   EVG   PPG
 2 1     Scot~ 31    RW    58    11      28      39      6     45    7     2
 3 2     Mari~ 18    LW    71    18      18      36      -6    32    12    6
 4 3     Ľubo~ 32    D     80    11      23      34      -8    52    7     4
 5 4     Wes ~ 30    C     82    18      12      30      -8    37    11    0
 6 5     Fili~ 24    D     75    9       21      30      -6    28    5     4
 7 6     Darb~ 28    LW    72    18      11      29      1     36    14    3
 8 7     Jim ~ 32    C     68    7       22      29      -6    80    7     0
 9 8     Antt~ 27    LW    82    12      16      28      -7    24    10    0
10 9     Stac~ 26    C     76    7       20      27      3     20    6     1
# i 30 more rows
# i 10 more variables: Goals <chr>, Goals <chr>, Assists <chr>, Assists <chr>,
#   Assists <chr>, Shots <chr>, Shots <chr>, `Ice Time` <chr>,
#   `Ice Time` <chr>, `` <chr>
```

3) Place the code from (2) into a function where the user can input a team and year. You would then adjust the url accordingly and produce a clean table for the user.

```r
hockey_stats <- function(team, year){
  url <- str_c("https://www.hockey-reference.com/teams/", team, "/", year, ".html")
  session <- bow(url, force = TRUE)
  result <- scrape(session) |>
    html_nodes(css = "table") |>
    html_table(header = TRUE, fill = TRUE)
  player_tibble <- result[[4]] |>
    row_to_names(row_number = 1) |>
    clean_names() |>
    mutate(year = year) |>
    select(player, year, age, pos, gp, pts)
  player_tibble
}

hockey_stats("MIN", 2020)
```

```
No encoding supplied: defaulting to UTF-8.
```

```
# A tibble: 31 x 6
   player             year age   pos   gp    pts  
   <chr>             <dbl> <chr> <chr> <chr> <chr>
 1 Kevin Fiala        2020 23    LW    64    54   
 2 Ryan Suter         2020 35    D     69    48   
 3 Eric Staal         2020 35    C     66    47   
 4 Zach Parise        2020 35    LW    69    46   
 5 Mats Zuccarello    2020 32    LW    65    37   
 6 Jared Spurgeon     2020 30    D     62    32   
 7 Luke Kunin         2020 22    C     63    31   
 8 Jason Zucker       2020 28    LW    45    29   
 9 Joel Eriksson Ek   2020 23    C     62    29   
10 Jordan Greenway    2020 22    LW    67    28   
# i 21 more rows
```

4) Use `map2` and `list_rbind` to build one data set containing Minnesota Wild data from 2001-2004.

```
teams <- rep("MIN", 4)
years <- 2001:2004
temp <- map2(teams, years, hockey_stats)
```

```
No encoding supplied: defaulting to UTF-8.
No encoding supplied: defaulting to UTF-8.
No encoding supplied: defaulting to UTF-8.
No encoding supplied: defaulting to UTF-8.
```

```
hockey_data_4yrs <- list_rbind(temp)
hockey_data_4yrs
```

```
# A tibble: 141 x 6
   player              year age   pos   gp    pts  
   <chr>              <int> <chr> <chr> <chr> <chr>
 1 Scott Pellerin      2001 31    RW    58    39   
 2 Marián Gáborík      2001 18    LW    71    36   
 3 Ľubomír Sekeráš     2001 32    D     80    34   
 4 Wes Walz            2001 30    C     82    30   
 5 Filip Kuba          2001 24    D     75    30   
 6 Darby Hendrickson   2001 28    LW    72    29   
 7 Jim Dowd            2001 32    C     68    29   
 8 Antti Laaksonen     2001 27    LW    82    28   
 9 Stacy Roest         2001 26    C     76    27   
10 Aaron Gavey         2001 26    C     75    24   
# i 131 more rows
```

09 Pause to Ponder

[Pause to Ponder:] Create a function to scrape a single NIH press release page by filling missing pieces labeled `???`:

```r
# Helper function to reduce html_nodes() |> html_text() code duplication
get_text_from_page <- function(page, css_selector) {
  page |>
  html_nodes(css_selector) |>
  html_text()
}

# Main function to scrape and tidy desired attributes
scrape_page <- function(url) {
    Sys.sleep(2)
    page <- read_html(url)
    article_titles <- get_text_from_page(page, ".teaser-title")
    article_dates <- get_text_from_page(page, ".date-display-single")
    article_dates <- mdy(article_dates)
    article_description <- get_text_from_page(page, ".teaser-description")
    article_description <- str_trim(str_replace(article_description,
                                                ".*\\n",
                                                "")
                                    )

    tibble(
        article_titles, article_dates, article_description
    )
}

scrape_page("https://www.nih.gov/news-events/news-releases")
```

```
# A tibble: 10 x 3
   article_titles                                 article_dates article_description
   <chr>                                          <date>        <chr>
 1 Kidney transplantation between donors and ~ 2024-10-16    NIH-funded study p~
 2 Mpox vaccine is safe and generates a robus~ 2024-10-16    NIH clinical trial~
 3 NIH and FDA leaders call for innovation in~ 2024-10-15    Commentary emphasi~
 4 Alzheimer's disease may damage the brain i~ 2024-10-15    NIH-funded brain m~
 5 First wave of COVID-19 increased risk of h~ 2024-10-10    NIH-funded study f~
 6 NIH launches large study to tackle type 2 ~ 2024-10-09    Effort to identify~
 7 Researchers fully map neural connections o~ 2024-10-02    NIH-supported mile~
 8 Scientists discover gene responsible for r~ 2024-09-26    NIH-supported find~
 9 Statement by NIH on Research Misconduct Fi~ 2024-09-26    NIH takes very ser~
10 Higher doses of buprenorphine may improve ~ 2024-09-25    NIH-funded analysi~
```

[**Pause to Ponder:**] Use a for loop over the first 5 pages:

```
pages <- vector("list", length = 5)

for (i in 1:5) {
    base_url <- "https://www.nih.gov/news-events/news-releases"
    if (i==1) {
        url <- base_url
    } else {
        url <- str_c(base_url, "?page=", i-1)
    }
    pages[[i]] <- scrape_page(url)
}

df_articles <- bind_rows(pages)
head(df_articles)
```

```
# A tibble: 6 x 3
  article_titles                                  article_dates article_description
  <chr>                                           <date>        <chr>
1 Kidney transplantation between donors and r~ 2024-10-16    NIH-funded study p~
```

```
2 Mpox vaccine is safe and generates a robust~ 2024-10-16    NIH clinical trial~
3 NIH and FDA leaders call for innovation in ~ 2024-10-15    Commentary emphasi~
4 Alzheimer's disease may damage the brain in~ 2024-10-15    NIH-funded brain m~
5 First wave of COVID-19 increased risk of he~ 2024-10-10    NIH-funded study f~
6 NIH launches large study to tackle type 2 d~ 2024-10-09    Effort to identify~
```

[**Pause to Ponder:**] Use map functions in the purrr package:

```r
base_url <- "https://www.nih.gov/news-events/news-releases"
urls_all_pages <- c(base_url, str_c(base_url, "?page=", as.character(1:4)))

pages2 <- purrr::map(urls_all_pages, scrape_page)
df_articles2 <- bind_rows(pages2)
head(df_articles2)
```

```
# A tibble: 6 x 3
  article_titles                               article_dates article_description
  <chr>                                        <date>        <chr>
1 Kidney transplantation between donors and r~ 2024-10-16    NIH-funded study p~
2 Mpox vaccine is safe and generates a robust~ 2024-10-16    NIH clinical trial~
3 NIH and FDA leaders call for innovation in ~ 2024-10-15    Commentary emphasi~
4 Alzheimer's disease may damage the brain in~ 2024-10-15    NIH-funded brain m~
5 First wave of COVID-19 increased risk of he~ 2024-10-10    NIH-funded study f~
6 NIH launches large study to tackle type 2 d~ 2024-10-09    Effort to identify~
```