HW #3

Gwynyn Hayes

07_apis.qmd: On Your Own #2-3

```
# function to allow user inputs

MN_tract_data <- function(year, county, variables) {
  tidycensus::get_acs(
    Sys.sleep(0.5),
    year = year,
    state = "MN",
    geography = "tract",
    variables = variables,
    output = "wide",
    geometry = TRUE,
    county = county
  ) |>
    mutate(year = year)
}

# Should really build in checks so that county is in MN, year is in
#   proper range, and variables are part of ACS1 data set

my_data <- MN_tract_data(year = 2021,
              county = "Hennepin",
              variables = c("B01003_001", "B19013_001"))
```

```
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
```

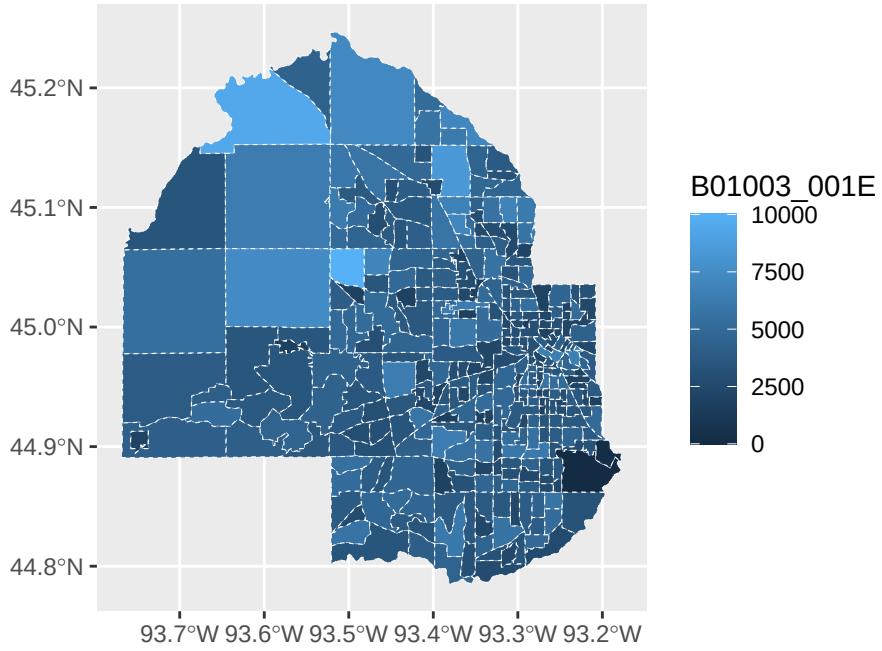
```
`census_api_key()` function to use it throughout your tidycensus session.
This warning is displayed once per session.

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

  |
  |                                                                      |   0%
  |
  |=                                                                     |   2%
  |
  |==                                                                    |   3%
  |
  |===                                                                   |   5%
  |
  |====                                                                  |   6%
  |
  |======                                                                |   8%
  |
  |=======                                                               |  10%
  |
  |========                                                              |  11%
  |
  |=========                                                             |  13%
  |
  |==========                                                            |  14%
  |
  |===========                                                           |  16%
  |
  |============                                                          |  18%
  |
  |==============                                                        |  19%
  |
  |===============                                                       |  21%
  |
  |================                                                      |  22%
  |
  |==================                                                    |  24%
  |
  |===================                                                   |  25%
  |
  |====================                                                  |  27%
```

```
|
|====================                                   |  28%
|
|====================                                   |  30%
|
|======================                                 |  32%
|
|========================                               |  33%
|
|========================                               |  35%
|
|==========================                             |  37%
|
|==========================                             |  38%
|
|============================                           |  40%
|
|============================                           |  41%
|
|==============================                         |  43%
|
|==============================                         |  45%
|
|================================                       |  46%
|
|================================                       |  48%
|
|==================================                     |  49%
|
|====================================                   |  52%
|
|====================================                   |  54%
|
|======================================                 |  56%
|
|======================================                 |  57%
|
|========================================               |  59%
|
|========================================               |  60%
|
|==========================================             |  62%
|
```

```
|=========================================                              |  64%
|
|==========================================                             |  65%
|
|===========================================                            |  67%
|
|============================================                           |  68%
|
|=============================================                          |  70%
|
|==============================================                         |  72%
|
|===============================================                        |  73%
|
|================================================                       |  75%
|
|==================================================                     |  78%
|
|===================================================                    |  79%
|
|====================================================                   |  81%
|
|======================================================                 |  83%
|
|=======================================================                |  84%
|
|=========================================================              |  86%
|
|==========================================================             |  88%
|
|============================================================           |  89%
|
|=============================================================          |  91%
|
|==============================================================         |  92%
|
|================================================================       |  94%
|
|=================================================================      |  96%
|
|==================================================================     |  97%
|
|====================================================================== |  99%
```

```
  |
  |======================================================================| 100%
```

```r
ggplot(data = my_data) +
  geom_sf(aes(fill = B01003_001E), colour = "white", linetype = 2)
```




```r
my_data <- MN_tract_data(year = 2022,
              county = "Rice",
              variables = c("B01003_001", "B19013_001"))
```

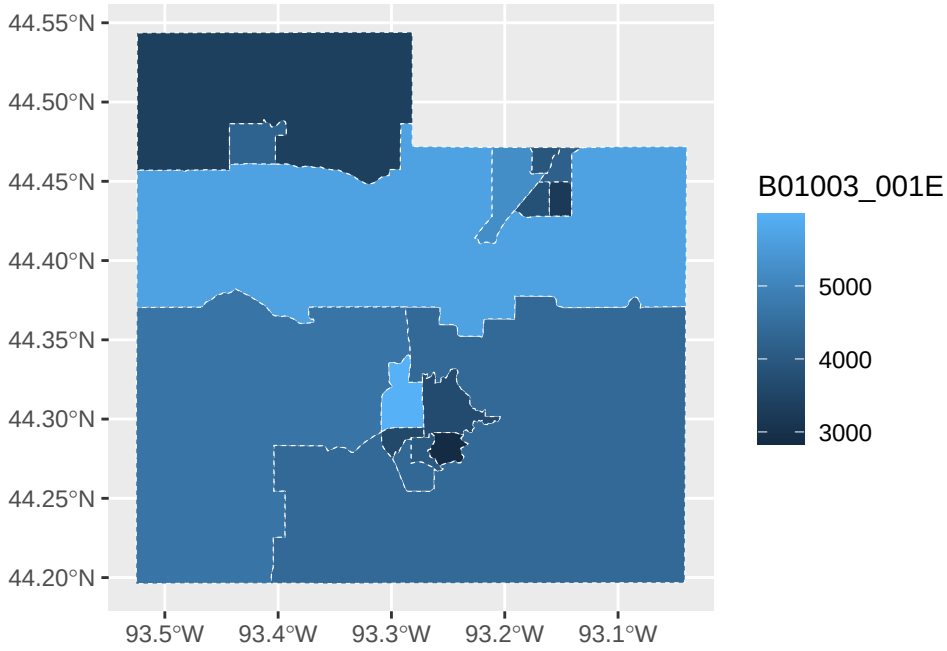
```
Getting data from the 2018-2022 5-year ACS

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


  |
  |                                                                      |   0%
  |
  |=                                                                     |   2%
  |
  |==                                                                    |   3%
```

```
|
|===                                                                   |   5%
|
|=====                                                                 |   7%
|
|======                                                                |   8%
|
|=======                                                               |  10%
|
|========                                                              |  11%
|
|=========                                                             |  13%
|
|==========                                                            |  15%
|
|===========                                                           |  16%
|
|=============                                                         |  18%
|
|==============                                                        |  20%
|
|================                                                      |  21%
|
|================                                                      |  23%
|
|==================                                                    |  24%
|
|====================                                                  |  27%
|
|====================                                                  |  29%
|
|======================                                                |  31%
|
|=======================                                               |  32%
|
|========================                                              |  34%
|
|=========================                                             |  35%
|
|==========================                                            |  37%
|
|============================                                          |  39%
|
```

```
|============================                                   |  40%
|
|=============================                                  |  42%
|
|==============================                                 |  44%
|
|===============================                                |  45%
|
|================================                               |  47%
|
|=================================                              |  49%
|
|==================================                             |  50%
|
|====================================                           |  53%
|
|=====================================                          |  55%
|
|======================================                         |  57%
|
|=======================================                        |  58%
|
|========================================                       |  60%
|
|=========================================                      |  61%
|
|==========================================                     |  63%
|
|===========================================                    |  65%
|
|============================================                   |  66%
|
|=============================================                  |  68%
|
|==============================================                 |  70%
|
|===============================================                |  71%
|
|================================================               |  73%
|
|=================================================              |  74%
|
|==================================================             |  76%
```

```
  |
  |=========================================================             |  79%
  |
  |==========================================================            |  81%
  |
  |===========================================================           |  82%
  |
  |============================================================          |  84%
  |
  |=============================================================         |  86%
  |
  |==============================================================        |  87%
  |
  |================================================================      |  89%
  |
  |=================================================================     |  91%
  |
  |==================================================================    |  92%
  |
  |====================================================================  |  94%
  |
  |===================================================================== |  95%
  |
  |======================================================================|  97%
  |
  |======================================================================| |  99%
  |
  |======================================================================| 100%
```

```
ggplot(data = my_data) +
  geom_sf(aes(fill = B01003_001E), colour = "white", linetype = 2)
```




```
# Try other variables:
#  - B25077_001 is median home price
#  - B02001_002 is number of white residents
#  - etc.
# alt
```

```
# To examine trends over time in Rice County
2019:2021 |>
  purrr::map(\(x)
    MN_tract_data(
      x,
      county = "Rice",
      variables = c("B01003_001", "B19013_001")
    )
  ) |>
  list_rbind()
```

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

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

```
|
|                                                                      |   0%
|
|=                                                                     |   2%
|
|===                                                                   |   4%
|
|====                                                                  |   5%
|
|=====                                                                 |   7%
|
|======                                                                |   9%
|
|========                                                              |  11%
|
|=========                                                             |  13%
|
|==========                                                            |  14%
|
|===========                                                           |  16%
|
|=============                                                         |  18%
|
|==============                                                        |  20%
|
|================                                                      |  22%
|
|=================                                                     |  23%
|
|==================                                                    |  25%
|
|====================                                                  |  27%
|
|=====================                                                 |  29%
|
|======================                                                |  30%
|
|========================                                              |  32%
|
|=========================                                             |  34%
|
|=========================                                             |  36%
|
```

```
|=========================                                              |  38%
|
|==========================                                             |  39%
|
|============================                                           |  41%
|
|============================                                           |  43%
|
|=============================                                          |  45%
|
|==============================                                         |  47%
|
|==============================                                         |  48%
|
|===============================                                        |  50%
|
|================================                                       |  52%
|
|=================================                                      |  54%
|
|=================================                                      |  56%
|
|==================================                                     |  57%
|
|==================================                                     |  59%
|
|===================================                                    |  61%
|
|===================================                                    |  63%
|
|====================================                                   |  64%
|
|====================================                                   |  66%
|
|=====================================                                  |  68%
|
|=====================================                                  |  70%
|
|======================================                                 |  72%
|
|======================================                                 |  73%
|
|=======================================                                |  75%
```

```
  |
  |=================================================                 |  77%
  |
  |==================================================                |  79%
  |
  |===================================================               |  81%
  |
  |====================================================              |  82%
  |
  |=====================================================             |  84%
  |
  |======================================================            |  86%
  |
  |=======================================================           |  88%
  |
  |========================================================          |  89%
  |
  |=========================================================         |  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

  |
  |                                                                   |   0%
  |
  |=                                                                  |   2%
  |
  |==                                                                 |   3%
  |
  |===                                                                |   5%
  |
```

```
|====                                                             |   6%
|
|=====                                                            |   8%
|
|======                                                           |   9%
|
|=======                                                          |  11%
|
|========                                                         |  12%
|
|==========                                                       |  14%
|
|===========                                                      |  15%
|
|============                                                     |  17%
|
|==============                                                   |  18%
|
|===============                                                  |  20%
|
|================                                                 |  21%
|
|================                                                 |  23%
|
|==================                                               |  24%
|
|===================                                              |  25%
|
|====================                                             |  27%
|
|=====================                                            |  28%
|
|======================                                           |  30%
|
|=======================                                          |  31%
|
|========================                                         |  33%
|
|=========================                                        |  34%
|
|==========================                                       |  36%
|
|===========================                                      |  38%
```

```
|
|============================                                     |  39%
|
|============================                                     |  41%
|
|=============================                                    |  42%
|
|==============================                                   |  44%
|
|==============================                                   |  45%
|
|===============================                                  |  47%
|
|================================                                 |  49%
|
|================================                                 |  51%
|
|=================================                                |  53%
|
|=================================                                |  54%
|
|==================================                               |  56%
|
|==================================                               |  57%
|
|===================================                              |  59%
|
|===================================                              |  60%
|
|====================================                             |  62%
|
|====================================                             |  63%
|
|=====================================                            |  65%
|
|=====================================                            |  66%
|
|======================================                           |  68%
|
|======================================                           |  69%
|
|=======================================                          |  71%
|
```

```
  |=================================================                    |  74%
  |
  |==================================================                   |  75%
  |
  |====================================================                 |  77%
  |
  |=====================================================                |  78%
  |
  |======================================================               |  80%
  |
  |=======================================================              |  81%
  |
  |=========================================================            |  83%
  |
  |==========================================================           |  84%
  |
  |============================================================         |  86%
  |
  |=============================================================        |  87%
  |
  |==============================================================       |  89%
  |
  |================================================================     |  90%
  |
  |==================================================================   |  92%
  |
  |===================================================================  |  93%
  |
  |===================================================================== |  95%
  |
  |======================================================================|  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
```

	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	year
1	273	63989	9273	MULTIPOLYGON (((-93.19137 4...	2019
2	168	85952	2758	MULTIPOLYGON (((-93.40564 4...	2019
3	190	78343	4242	MULTIPOLYGON (((-93.52521 4...	2019
4	245	92321	14200	MULTIPOLYGON (((-93.16075 4...	2019
5	333	50368	9979	MULTIPOLYGON (((-93.17615 4...	2019
6	465	48403	7679	MULTIPOLYGON (((-93.29819 4...	2019
7	456	44417	10552	MULTIPOLYGON (((-93.30904 4...	2019
8	414	67868	9422	MULTIPOLYGON (((-93.27265 4...	2019
9	326	91667	8106	MULTIPOLYGON (((-93.52452 4...	2019
10	310	64479	12376	MULTIPOLYGON (((-93.22644 4...	2019
11	177	101359	4104	MULTIPOLYGON (((-93.5246 44...	2019
12	410	45230	12887	MULTIPOLYGON (((-93.30888 4...	2019
13	270	66188	9179	MULTIPOLYGON (((-93.16981 4...	2019
14	316	64792	13256	MULTIPOLYGON (((-93.16981 4...	2020
15	536	63500	7351	MULTIPOLYGON (((-93.1909 44...	2020
16	703	67625	23325	MULTIPOLYGON (((-93.29829 4...	2020
17	473	104011	5648	MULTIPOLYGON (((-93.5246 44...	2020
18	218	73750	13139	MULTIPOLYGON (((-93.26704 4...	2020
19	296	86094	3438	MULTIPOLYGON (((-93.40564 4...	2020
20	244	79068	4902	MULTIPOLYGON (((-93.52518 4...	2020
21	525	52936	10436	MULTIPOLYGON (((-93.17615 4...	2020
22	199	96023	13649	MULTIPOLYGON (((-93.44292 4...	2020
23	437	63924	8715	MULTIPOLYGON (((-93.28272 4...	2020
24	425	49811	16864	MULTIPOLYGON (((-93.27265 4...	2020
25	566	51595	9615	MULTIPOLYGON (((-93.30904 4...	2020
26	341	100516	11630	MULTIPOLYGON (((-93.16075 4...	2020
27	440	46750	15457	MULTIPOLYGON (((-93.30888 4...	2020
28	295	100563	15809	MULTIPOLYGON (((-93.52452 4...	2020
29	377	62078	5270	MULTIPOLYGON (((-93.22644 4...	2020
30	421	47059	15456	MULTIPOLYGON (((-93.30888 4...	2021
31	435	56319	4333	MULTIPOLYGON (((-93.17615 4...	2021
32	321	105952	8429	MULTIPOLYGON (((-93.16075 4...	2021
33	409	57126	13968	MULTIPOLYGON (((-93.27265 4...	2021
34	714	47344	9579	MULTIPOLYGON (((-93.30904 4...	2021
35	622	61193	23977	MULTIPOLYGON (((-93.29829 4...	2021
36	380	79063	15272	MULTIPOLYGON (((-93.16981 4...	2021
37	296	83911	7244	MULTIPOLYGON (((-93.52522 4...	2021
38	520	111711	10313	MULTIPOLYGON (((-93.5246 44...	2021
39	274	90179	4919	MULTIPOLYGON (((-93.40564 4...	2021
40	358	82500	20934	MULTIPOLYGON (((-93.26775 4...	2021
41	537	67219	9805	MULTIPOLYGON (((-93.1909 44...	2021
42	270	108490	1768	MULTIPOLYGON (((-93.52452 4...	2021

```
43         462       63679       12261 MULTIPOLYGON (((-93.28274 4... 2021
44         199       85789       20094 MULTIPOLYGON (((-93.44292 4... 2021
45         359       63835        4805 MULTIPOLYGON (((-93.22644 4... 2021
```

```r
# Or a little more simply
2019:2021 |>
  purrr::map(MN_tract_data,
             county = "Rice",
             variables = c("B01003_001", "B19013_001")
            ) |>
  list_rbind()
```

```
Getting data from the 2015-2019 5-year ACS
Downloading feature geometry from the Census website.  To cache shapefiles for use in future

Getting data from the 2016-2020 5-year ACS

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

Getting data from the 2017-2021 5-year ACS

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

         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 year
1          273       63989        9273 MULTIPOLYGON (((-93.19137 4... 2019
2          168       85952        2758 MULTIPOLYGON (((-93.40564 4... 2019
3          190       78343        4242 MULTIPOLYGON (((-93.52521 4... 2019
4          245       92321       14200 MULTIPOLYGON (((-93.16075 4... 2019
5          333       50368        9979 MULTIPOLYGON (((-93.17615 4... 2019
6          465       48403        7679 MULTIPOLYGON (((-93.29819 4... 2019
7          456       44417       10552 MULTIPOLYGON (((-93.30904 4... 2019
8          414       67868        9422 MULTIPOLYGON (((-93.27265 4... 2019
9          326       91667        8106 MULTIPOLYGON (((-93.52452 4... 2019
10         310       64479       12376 MULTIPOLYGON (((-93.22644 4... 2019
11         177      101359        4104 MULTIPOLYGON (((-93.5246 44... 2019
12         410       45230       12887 MULTIPOLYGON (((-93.30888 4... 2019
13         270       66188        9179 MULTIPOLYGON (((-93.16981 4... 2019
```

```
14         316       64792       13256 MULTIPOLYGON (((-93.16981 4... 2020
15         536       63500        7351 MULTIPOLYGON (((-93.1909 44... 2020
16         703       67625       23325 MULTIPOLYGON (((-93.29829 4... 2020
17         473      104011        5648 MULTIPOLYGON (((-93.5246 44... 2020
18         218       73750       13139 MULTIPOLYGON (((-93.26704 4... 2020
19         296       86094        3438 MULTIPOLYGON (((-93.40564 4... 2020
20         244       79068        4902 MULTIPOLYGON (((-93.52518 4... 2020
21         525       52936       10436 MULTIPOLYGON (((-93.17615 4... 2020
22         199       96023       13649 MULTIPOLYGON (((-93.44292 4... 2020
23         437       63924        8715 MULTIPOLYGON (((-93.28272 4... 2020
24         425       49811       16864 MULTIPOLYGON (((-93.27265 4... 2020
25         566       51595        9615 MULTIPOLYGON (((-93.30904 4... 2020
26         341      100516       11630 MULTIPOLYGON (((-93.16075 4... 2020
27         440       46750       15457 MULTIPOLYGON (((-93.30888 4... 2020
28         295      100563       15809 MULTIPOLYGON (((-93.52452 4... 2020
29         377       62078        5270 MULTIPOLYGON (((-93.22644 4... 2020
30         421       47059       15456 MULTIPOLYGON (((-93.30888 4... 2021
31         435       56319        4333 MULTIPOLYGON (((-93.17615 4... 2021
32         321      105952        8429 MULTIPOLYGON (((-93.16075 4... 2021
33         409       57126       13968 MULTIPOLYGON (((-93.27265 4... 2021
34         714       47344        9579 MULTIPOLYGON (((-93.30904 4... 2021
35         622       61193       23977 MULTIPOLYGON (((-93.29829 4... 2021
36         380       79063       15272 MULTIPOLYGON (((-93.16981 4... 2021
37         296       83911        7244 MULTIPOLYGON (((-93.52522 4... 2021
38         520      111711       10313 MULTIPOLYGON (((-93.5246 44... 2021
39         274       90179        4919 MULTIPOLYGON (((-93.40564 4... 2021
40         358       82500       20934 MULTIPOLYGON (((-93.26775 4... 2021
41         537       67219        9805 MULTIPOLYGON (((-93.1909 44... 2021
42         270      108490        1768 MULTIPOLYGON (((-93.52452 4... 2021
43         462       63679       12261 MULTIPOLYGON (((-93.28274 4... 2021
44         199       85789       20094 MULTIPOLYGON (((-93.44292 4... 2021
45         359       63835        4805 MULTIPOLYGON (((-93.22644 4... 2021
```

OMDB

```r
# I used the first line to store my OMDB API key in .Renviron
# Sys.setenv(OMDB_KEY = "98cc43c7")
myapikey <- Sys.getenv("OMDB_KEY")

# Find url exploring examples at omdbapi.com
url <- str_c("http://www.omdbapi.com/?t=Coco&y=2017&apikey=", myapikey)
```

```
coco <- GET(url)   # coco holds response from server
coco               # Status of 200 is good!

details <- content(coco, "parse")
details                          # get a list of 25 pieces of information
details$Year                     # how to access details
details[[2]]
```

```
# Must figure out pattern in URL for obtaining different movies
#  - try searching for others
movies <- c("The+Pirate+Fairy", "Knives+Out", "Fighting+with+my+Family", "The+Princess+Diarie

# Set up empty tibble
omdb <- tibble(Title = character(), Language = character(), Writer = character(),
       Awards = character(), Plot = character())

# 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=", myapikey)
  Sys.sleep(0.5)
  onemovie <- GET(url)
  detail <- content(onemovie, "parse")
  omdb[i,1] <- detail$Title
  omdb[i,2] <- detail$Language
  omdb[i,3] <- detail$Writer
  omdb[i,4] <- detail$Awards
  omdb[i,5] <- detail$Plot
}

omdb

#  could use stringr functions to further organize this data - separate
#    different genres, different actors, etc.
```

08_table_scraping.qmd: On Your Own #2.2-2.4

2. We would like to create a tibble with 4 years of data (2001-2004) from the Minnesota Wild hockey team. Specifically, we are interested in the "Scoring Regular Season" table from this webpage: https://www.hockey-reference.com/teams/MIN/2001.html and the

similar webpages from 2002, 2003, and 2004. Your final tibble should have 6 columns: player, year, age, pos (position), gp (games played), and pts (points).

You should (a) write a function called `hockey_stats` with inputs for team and year to scrape data from the "scoring Regular Season" table, and (b) use iteration techniques to scrape and combine 4 years worth of data. Here are some functions you might consider:

- `row_to_names(row_number = 1)` from the `janitor` package
- `clean_names()` also from the `janitor` package
- `bow()` and `scrape()` from the `polite` package
- `str_c()` from the `stringr` package (for creating urls with user inputs)
- `map2()` and `list_rbind()` for iterating and combining years

Try following these steps:

[SKIP] 1) Be sure you can find and clean the correct table from the 2021 season.

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

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.

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

```
hockey_stats <- function(team, year) {
  # Build the URL
  base_url <- str_c("https://www.hockey-reference.com/teams/", team, "/", year, ".html")

  session <- bow(base_url)
  page <- scrape(session)

  table_node <- html_node(page, css = "table")


  table <- html_table(table_node, header = TRUE, fill = TRUE)[[4]]

  table_clean <- table |>
    select(player, age, pos, gp, pts) |>
    mutate(year = year) |>
    relocate(year, .before = player) |>
    filter(!is.na(player))

  return(table_clean)
}
```

“‘{r years <- 2001:2004 team <- “MIN”

all_data <- map2(team, years, hockey_stats) |> list_rbind()

print(all_data)

09_web_scraping.qmd Pause to Ponder - 3 items on NIH News Releases right before the On Your

::: {.cell}

```{.r .cell-code}
# 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(
      title = article_titles,
      pub_date = article_dates,
      description = article_description
    )
}
```

:::

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

** Said it could find function even though it runs and works properly “‘{r pages <- vector(“list”, length = 6) pos <- 0

for (i in 2025:2024) { for (j in 0:2) { pos <- pos + 1 url <- str_c(“https://www.nih.gov/news-events/news-releases?”, i, “&page=”, j, “&1=”) pages[[pos]] <- scrape_page(url) } }

df_articles <- bind_rows(pages) head(df_articles)

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

::: {.cell}

```{.r .cell-code}
base_url <- "https://www.nih.gov/news-events/news-releases?page="
urls_all_pages <- c(base_url, str_c(base_url, 1:5))

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

:::