Final Project JOUR472

This summer, I???ll be interning with The Baltimore Sun and covering Carroll and Howard counties. Because of how voting, demographics and other information in various counties have changed over time, I want to explore these fields and understand the counties I will soon be reporting on. I hope to use this analysis to inform my reporting this summer. The data I am using is from the U.S. Census and Maryland election website.

First, we need to install our packages.

```
library(tidyverse)
library(tidycensus)
library(rvest)
library(janitor)
library(formattable)
library(tinytex)
```

Now, to answer some of my questions about Howard and Carroll county demographics, I need to grab Census data. I'll take it with my API key.

```
[1] "423292f0a0d4b5f0fe4ce722f283f6af07c21c84"
```

First, I wanted to understand the demographics of both Carroll and Howard counties and how that has changed over time. For this, I took U.S. Census Data from 2020 and 2010 to see how these populations have changed. With the rising cost of living as well, I would like to see how the diversity of these counties have changed. **Q** *How has diversity changed over time in both Howard and Carroll counties? What race of people populates each county the most?*

```
#load in census data through my API
acs_vars <- load_variables("acs5", year = 2020)

#grab maryland section, divide by county by race for both 2010 and 2020. I want to get the
x2020_md <- get_acs(geography = "county",
              variables = c(white = "B02001_002",
```

```
                            black = "B02001_003",
                            american_ind = "B02001_004",
                            asian = "B02001_005",
                            pacif_islander = "B02001_006",
                            hispanic = "B03003_003"),
              state = "MD",
              year = 2020)

x2010_md <- get_acs(geography = "county",
              variables = c(white = "B02001_002",
                            black = "B02001_003",
                            american_ind = "B02001_004",
                            asian = "B02001_005",
                            pacif_islander = "B02001_006",
                            hispanic = "B03003_003"),
              state = "MD",
              year = 2010)
```

In these two dataframes, I now have the GEOID, name, variable, estimate and margin of error. This is all in regards to racial population, and variable determines which group is being talked about – it's divided by race, and helps me sort out the populations. GEOID is the geographical identification of the area in NAME, which is just the name of the county. I like to combine dataframes by GEOID because it's very precise and most often doesn't have any spelling errors in the data, so has an easy time to join. Estimate is just that – the estimate of the number of people in the specific population. Margin of error is how many people could be given or taken from the estimate, just essentially the wiggle room in the data the Census has that shows where things could possibly go wrong.

Now, to join the dataframes and filter the counties out that I don't want.

```
#now want to filter out the counties! but first, i'll join the two dataframes. as mentione
md_total <- x2020_md %>%
  inner_join(x2010_md, by = c("GEOID", "variable")) %>%
  rename(estimate_2020 = estimate.x,
         estimate_2010 = estimate.y,
         name = "NAME.x") %>%
  select(GEOID, name, variable, estimate_2020, estimate_2010)

#filter out other counties and mutate for pct_change and pct_whole, then i just am going t
demo_howard <- md_total %>%
  filter(name == "Howard County, Maryland") %>%
  mutate(howard_pct_change = (((estimate_2020-estimate_2010)/estimate_2010)*100)) %>%
```

```
  mutate(all_2020 = sum(estimate_2020)) %>%
  mutate(all_2010 = sum(estimate_2010)) %>%
  mutate(howard_pct_whole_20 = percent(estimate_2020/all_2020)) %>%
  mutate(howard_pct_whole_10 = percent(estimate_2010/all_2010))

demo_carroll <- md_total %>%
  filter(name == "Carroll County, Maryland") %>%
  mutate(carroll_pct_change = (((estimate_2020-estimate_2010)/estimate_2010)*100)) %>%
  mutate(all_2020 = sum(estimate_2020)) %>%
  mutate(all_2010 = sum(estimate_2010)) %>%
  mutate(carroll_pct_whole_20 = percent(estimate_2020/all_2020)) %>%
  mutate(carroll_pct_whole_10 = percent(estimate_2010/all_2010))

#now want to view both of my dataframes -- i liked having two separate ones since it was e
view(demo_carroll)

view(demo_howard)

#based on this information, it appears that both Howard County and Carroll County have bec

#In Howard County, the white population decreased from about 33% in 2010 to about 28% of t

#In Carroll County, the white population decreased from an about 47% of the whole populati
```

Now that we've grabbed this information for Howard and Carroll counties, I want to understand how this compares to the demographic breakdown in Maryland and how that has changed. So now, let's do the same thing we did above for the entire state of Maryland.

```
#maryland 2020 and maryland 2010 demographics -- i can pull from the same parts of the Cen
x2020_state <- get_acs(geography = "state",
              variables = c(white = "B02001_002",
                            black = "B02001_003",
                            american_ind = "B02001_004",
                            asian = "B02001_005",
                            pacif_islander = "B02001_006",
                            hispanic = "B03003_003"),
              state = "MD",
              year = 2020)

x2010_state <- get_acs(geography = "state",
              variables = c(white = "B02001_002",
```

```
                            black = "B02001_003",
                            american_ind = "B02001_004",
                            asian = "B02001_005",
                            pacif_islander = "B02001_006",
                            hispanic = "B03003_003"),
              state = "MD",
              year = 2010)

head(x2010_state)
```

```
# A tibble: 6 x 5
  GEOID NAME     variable       estimate   moe
  <chr> <chr>    <chr>             <dbl> <dbl>
1 24    Maryland white           3396216  4623
2 24    Maryland black           1665235  3361
3 24    Maryland american_ind      16213  1315
4 24    Maryland asian            304574  1595
5 24    Maryland pacif_islander     2977   587
6 24    Maryland hispanic         429946    NA
```

```
head(x2020_state)
```

```
# A tibble: 6 x 5
  GEOID NAME     variable       estimate   moe
  <chr> <chr>    <chr>             <dbl> <dbl>
1 24    Maryland white           3275048  6375
2 24    Maryland black           1803128  5297
3 24    Maryland american_ind      15860  1005
4 24    Maryland asian            384429  2538
5 24    Maryland pacif_islander     2650   417
6 24    Maryland hispanic         619418    NA
```

The values I have in these dataframes are the same – GEOID, name, variable, estimate and margin of error. Since I'm pulling from the same place, these also have the same meanings, as described above.

Let's join the 2020 and 2010 dataframes to compare the two census information bits as above.

```
#joining dataframes, mutating to make new columns -- going to make pct_change to see how p
state_total <- x2020_state %>%
```

```
  inner_join(x2010_state, by = c("GEOID", "variable")) %>%
  rename(estimate_2020 = estimate.x,
         estimate_2010 = estimate.y,
         name = "NAME.x") %>%
  mutate(state_pct_change = (percent((estimate_2020-estimate_2010)/estimate_2010))) %>%
  mutate(total_2010 = sum(estimate_2010)) %>%
  mutate(total_2020 = sum(estimate_2020)) %>%
  mutate(state_pct_total_2020 = percent(estimate_2020/total_2020)) %>%
  mutate(state_pct_total_2010 = percent(estimate_2010/total_2010)) %>%
  select(GEOID, name, variable, estimate_2020, total_2020,state_pct_total_2020,estimate_20

view(state_total)

#The white population in Maryland out of a percentage of the whole changed from about 30%
```

Finally on this analysis, I'm going to merge Carroll County demographics, Howard County demographics and Maryland demographics into one big dataframe. I'll do this so that way I can make a graphic to demonstrate the changes in the data.

```
#joining the state with carroll and cleaning up the names
state_carroll <- demo_carroll %>%
  inner_join(state_total, by = "variable") %>%
  rename(x20_carroll_estimate = "estimate_2020.x",
         x10_carroll_estimate = "estimate_2010.x",
         x20_state_estimate = "estimate_2020.y",
         x10_state_estimate = "estimate_2010.y"
         ) %>%
  select(variable, x20_carroll_estimate, x10_carroll_estimate, carroll_pct_change, x20_sta
```

Now, let's join the new state_carroll dataframe with the howard county dataframe.

```
state_howard_carroll <- state_carroll %>%
  inner_join(demo_howard, by = "variable") %>%
  rename(x20_howard_estimate = "estimate_2020",
         x10_howard_estimate = "estimate_2010") %>%
  select(variable, x20_carroll_estimate, x10_carroll_estimate, carroll_pct_change, x20_how

head(state_howard_carroll)
```

```
# A tibble: 6 x 10
  variable       x20_carroll_estimate x10_carroll_estimate carroll_pct_change
  <chr>                         <dbl>                <dbl>              <dbl>
```

```
1 white                          152180              155954              -2.42
2 black                            6088                5461              11.5
3 american_ind                      437                 341              28.2
4 asian                            3437                2291              50.0
5 pacif_islander                     73                   0              Inf
6 hispanic                         6217                4085              52.2
# i 6 more variables: x20_howard_estimate <dbl>, x10_howard_estimate <dbl>,
#   howard_pct_change <dbl>, x20_state_estimate <dbl>,
#   x10_state_estimate <dbl>, state_pct_change <formttbl>
```

Here's a graphic visualizing this data:

```
cat('<div class="flourish-embed flourish-chart" data-src="visualisation/17869350"><script
```

I also wanted to see how voting trends have changed in Carroll County. *How does Carroll County lean in voting Republican vs. Democratic? Has this changed in the past few election cycles?*

To answer this question, I am scraping this website for each election year in Carroll County and the entire state of Maryland. In it, there are quite a few variables – name, party, early voting, election day, absentee/provisional, total and percentage. Name is the title of the president and vice president running for office, and party is the party affiliation of the candidate. Each other column breaks down the number of people that voted in a variety of ways, including absentee, mail-in, early voting or on election day. The total column just totals all of the people who voted for that candidate, and the percent column describes what percentage of the whole voted for that particular candidate. As shown later down, I ended up doing a lot of cleaning on Carroll County's information to ensure the columns were how I wanted them to be named.

```
#read in the html and extract all the tables
x20_results <- "https://elections.maryland.gov/elections/2020/results/general/gen_results_
  read_html() %>%
  html_table()

x20_results <- x20_results[[2]]

x16_results <- "https://elections.maryland.gov/elections/2016/results/general/gen_results_
  html_table()

x16_results <- x16_results[[1]]

x12_results <- "https://elections.maryland.gov/elections/2012/results/general/gen_results_
  read_html() %>%
```

```
  html_table()

x12_results <- x12_results[[1]]
```

Now, I'm going to clean my dataframes.

```
#had to do a bit more cleaning on this table -- for some reason, it recognized information
x20_results_clean <- x20_results %>%
  mutate(
    Early_Voting = as.numeric(gsub("[^0-9.]", "", `Early Voting`)),  #making things numeri
    Election_Day = as.numeric(gsub("[^0-9.]", "", `Election Day`)),
    By_Mail = as.numeric(gsub("[^0-9.]", "", `By Mail`)),
    Total = as.numeric(gsub("[^0-9.]", "", Total)),
    Percent = as.numeric(gsub("[^0-9.]", "", Percent))
  ) %>%
  clean_names()

x16_results_clean <- x16_results %>%
  mutate(
    Early_Voting = as.numeric(gsub("[^0-9.]", "", `Early Voting`)),
    Election_Day = as.numeric(gsub("[^0-9.]", "", `Election Day`)),
    Absentee_Provisional = as.numeric(gsub("[^0-9.]", "", `Absentee / Provisional`)),
    Total = as.numeric(gsub("[^0-9.]", "", Total)),
    Percentage = as.numeric(gsub("[^0-9.]", "", Percentage))
  ) %>%
  clean_names()

x12_results_clean <- x12_results %>%
  mutate(
    Early_Voting = as.numeric(gsub("[^0-9.]", "", `Early Voting`)),
    Election_Day = as.numeric(gsub("[^0-9.]", "", `Election Day`)),
    Absentee_Provisional = as.numeric(gsub("[^0-9.]", "", `Absentee / Provisional`)),
    Total = as.numeric(gsub("[^0-9.]", "", Total)),
    Percentage = as.numeric(gsub("[^0-9.]", "", Percentage))
  ) %>%
  clean_names()
```

Now, let's analyze the data we've now cleaned to find how Carroll County has leaned in since the 2012 election.

```r
x20_gen_results <- x20_results_clean %>%
  select(name, party, early_voting_2, election_day_2,by_mail_2,total,percent) %>%
  rename(early_voting = "early_voting_2",
         election_day = "election_day_2",
         by_mail = "by_mail_2")

x16_gen_results <- x16_results_clean %>%
  select(name, party, early_voting_2, election_day_2,absentee_provisional_2,total,percenta
  rename(early_voting = "early_voting_2",
         election_day = "election_day_2",
         absentee_provisional = "absentee_provisional_2",
         percent = "percentage")

x12_gen_results <- x12_results_clean %>%
  select(name, party, early_voting_2, election_day_2, absentee_provisional_2, total, perce
  rename(early_voting = "early_voting_2",
         election_day = "election_day_2",
         absentee_provisional = "absentee_provisional_2",
         percent = "percentage")

#Based on this analysis with percentage, it appears that Carroll County has always skewed

x20_gen_results
```

```
# A tibble: 32 x 7
   name                    party early_voting election_day by_mail total percent
   <chr>                   <chr>        <dbl>        <dbl>   <dbl> <dbl>   <dbl>
 1 Donald J. Trump and Mi~ Repu~        25406        21755   12658 60218    60
 2 Joe  Biden and Kamala ~ Demo~         8898         4648   22754 36456    36.3
 3 Jo  Jorgensen and Jere~ Libe~          545          714     754  2028     2
 4 Howie Gresham Hawkins ~ Green          175          166     195   537     0.5
 5 Jerome M. Segal and Jo~ Brea~           70           47      52   169     0.2
 6 Sharon  Wallace and Ka~ Demo~            0            0       0     0     0
 7 Dennis Andrew Ball (Wr~ Other            0            0       0     0     0
 8 Barbara  Bellar (Write~ Other            0            1       0     1     0
 9 President  Boddie (Wri~ Other            1            0       0     1     0
10 Mary Ruth Caro Simmons~ Other            0            0       0     0     0
# i 22 more rows
```

```r
x16_gen_results
```

```
# A tibble: 57 x 7
   name       party early_voting election_day absentee_provisional total percent
   <chr>      <chr>        <dbl>        <dbl>                <dbl> <dbl>   <dbl>
 1 Trump/ Pe~ Repu~        11477        44165                 2573 58215    63.4
 2 Clinton/ ~ Demo~         7110        17205                 2252 26567    28.9
 3 Johnson/ ~ Libe~          618         3394                  274  4286     4.7
 4 Stein/ Ba~ Green          143          818                   99  1060     1.2
 5 Paij  Bor~ Repu~            0            0                    0     0     0
 6 Joann  Br~ Repu~            0            1                    0     1     0
 7 Jacquelin~ Repu~            0            4                    0     4     0
 8 Stephen  ~ Repu~            0            0                    0     0     0
 9 Janet L. ~ Repu~            0            0                    0     0     0
10 Paul  Ada~ Demo~            1            0                    0     1     0
# i 47 more rows
```

```
x12_gen_results
```

```
# A tibble: 37 x 7
   name       party early_voting election_day absentee_provisional total percent
   <chr>      <chr>        <dbl>        <dbl>                <dbl> <dbl>   <dbl>
 1 Obama/Bid~ Demo~         3978        22089                 1872 27939    31.9
 2 Romney/Ry~ Repu~         6176        47784                 2801 56761    64.8
 3 Johnson/G~ Libe~          123         1393                   98  1614     1.8
 4 Stein/Hon~ Green           51          577                   46   674     0.8
 5 Briscoe/O~ Demo~            0            1                    0     1     0
 6 Dennis  K~ Demo~            0            2                    0     2     0
 7 Matthew  ~ Repu~            0            0                    0     0     0
 8 Barbara A~ Repu~            0            0                    0     0     0
 9 Santa  Cl~ Inde~            1           29                    0    30     0
10 Richard  ~ Inde~            0            0                    0     0     0
# i 27 more rows
```

I also want to compare this to generally how the entire state of Maryland has leaned in general elections. Does the trend of Carroll County lean in the same way as the entire state?

I scraped the same website to answer this question and the variables are the same – name, party, early voting, election day, by mail/absentee/provisional, total and percent. The numbers aren't the same but count the same things.

```
#pull in data from the state elections
x12_result_md <- "https://elections.maryland.gov/elections/2012/results/general/gen_result
  read_html() %>%
```

```
  html_table()

x12_result_md <- x12_result_md[[1]]

x16_result_md <- "https://elections.maryland.gov/elections/2016/results/general/gen_result
  read_html() %>%
  html_table()

x16_result_md <- x16_result_md[[1]]

x20_result_md <- "https://elections.maryland.gov/elections/2020/results/general/gen_result
  read_html() %>%
  html_table()

x20_result_md <- x20_result_md[[2]]
```

Now, let's see how Marylanders voted in the past three elections.

```
head(x12_result_md)
```

```
# A tibble: 6 x 7
  Name          Party `Early Voting` `Election Day` Absentee / Provision~1 Total
  <chr>         <chr> <chr>          <chr>          <chr>                  <chr>
1 Obama/ Biden  Demo~ 310,922        1,218,709      148,213                1,67~
2 Romney/ Ryan  Repu~ 113,077        792,564        66,228                 971,~
3 Johnson/ Gray Libe~ 2,447          25,303         2,445                  30,1~
4 Stein/ Honka~ Green 1,759          13,817         1,534                  17,1~
5 Briscoe/ Ogl~ Demo~ 5              13             0                      18
6 Dennis  Knil~ Demo~ 0              2              0                      2
# i abbreviated name: 1: `Absentee / Provisional`
# i 1 more variable: Percentage <chr>
```

```
head(x16_result_md)
```

```
# A tibble: 6 x 7
  Name          Party `Early Voting` `Election Day` Absentee / Provision~1 Total
  <chr>         <chr> <chr>          <chr>          <chr>                  <chr>
1 Trump/ Pence  Repu~ 229,827        654,557        58,785                 943,~
2 Clinton/ Kai~ Demo~ 616,187        897,159        164,582                1,67~
3 Johnson/ Weld Libe~ 16,049         55,938         7,618                  79,6~
```

```
4 Stein/ Baraka Green 8,056          24,221         3,668                  35,9~
5 Paij  Boring~ Repu~ 17             34             2                      53
6 Joann  Breiv~ Repu~ 5              15             0                      20
# i abbreviated name: 1: `Absentee / Provisional`
# i 1 more variable: Percentage <chr>
```

```
head(x20_result_md)
```

```
# A tibble: 6 x 8
  Name      Party `Early Voting` `Election Day` `By Mail` Prov. Total Percentage
  <chr>     <chr> <chr>          <chr>          <chr>     <chr> <chr> <chr>
1 Donald J~ Repu~ 453,227        249,123        248,240   25,8~ 976,~ 32.2%
2 Joe  Bid~ Demo~ 507,955        170,767        1,221,847 84,4~ 1,98~ 65.4%
3 Jo  Jorg~ Libe~ 9,735          8,447          14,242    1,064 33,4~ 1.1%
4 Howie Gr~ Green 5,624          3,769          5,711     695   15,7~ 0.5%
5 Jerome M~ Brea~ 2,528          1,500          1,577     279   5,884 0.2%
6 Sharon  ~ Demo~ 0              0              0         0     0     0.0%
```

```
#Based on these results, the percentage of voters in the state seems to have changed in re

#When we compare this to above in the Carroll County sections, it appears that support for
```

Now, to make a graphic on these trends. The numbers are nice, but I'm not sure I can understand what I'm reporting on until I see the data in a graphic format.

Graphic on Carroll County votes in general elections

```
cat('<iframe src="<div class="flourish-embed flourish-chart" data-src="visualisation/17783
```

<iframe src="

">

Graphic about people in Maryland in favor of Republican candidates in the 2020 election

```
cat('<iframe title="13 counties in Maryland had about 50% of voters in favor of Republican
</script>')
```

Lastly, I wanted to explore how the household median income changed in Carroll County and Howard County. I'll be using data from the U.S. Census. Similar to my first question, the

data is grabbed using my census API and pulled from the data that the census has. *How has the median household income changed in Carroll County and Howard County?*

```
#grab the data by county for median income
x20_md_median_income <- get_acs(geography = "county",
              variables = c(median_income = "B19013_001"),
              state = "MD",
              year = 2020)

x10_md_median_income <- get_acs(geography = "county",
              variables = c(median_income = "B19013_001"),
              state = "MD",
              year = 2010)
```

There are GEOID, median income, name, estimate and margin of error columns. The estimate is of the household median income by county, and margin of error describes the margin in which the estimate could be wrong, give or take that number. The variables included are by state and county, and I want to specifically grab Carroll and Howard Counties.

Now, to join and filter the dataframes.

```
#clean the data! I want to filter out for the specific counties I want and do some cleanig
md_median_income_total <- x20_md_median_income %>%
  inner_join(x10_md_median_income, by="GEOID", "variable") %>%
  rename(name = "NAME.x",
         x20_estimate = "estimate.x",
         x10_estimate = "estimate.y",
         variable = "variable.x")
```

Lastly, I'm going to use my newly combined dataframes to find the median income for each county.

```
howard_med_income <- md_median_income_total %>%
  filter(name == "Howard County, Maryland") %>%
  select(name, x20_estimate, x10_estimate, variable) %>%
  mutate(pct_change = (((x20_estimate - x10_estimate)/x10_estimate)*100)) %>%
  arrange(desc(pct_change))

carroll_med_income <- md_median_income_total %>%
  filter(name == "Carroll County, Maryland") %>%
  select(name, x20_estimate, x10_estimate, variable) %>%
  mutate(pct_change = (((x20_estimate - x10_estimate)/x10_estimate)*100)) %>%
```

```
  arrange(desc(pct_change))

head(howard_med_income)
```

```
# A tibble: 1 x 5
  name                    x20_estimate x10_estimate variable      pct_change
  <chr>                          <dbl>        <dbl> <chr>              <dbl>
1 Howard County, Maryland       124042       103273 median_income       20.1
```

```
head(carroll_med_income)
```

```
# A tibble: 1 x 5
  name                     x20_estimate x10_estimate variable      pct_change
  <chr>                           <dbl>        <dbl> <chr>              <dbl>
1 Carroll County, Maryland        99569        81621 median_income       22.0
```

```
#The household median income in Carroll County increased by about 22%, according to Census
```

I also want to compare this to trends in the state to see if Carroll and Howard counties follow state trends with a rising median income.

```
#pull census data
x20_state_median_income <- get_acs(geography = "state",
              variables = c(median_income = "B19013_001"),
              state = "MD",
              year = 2020)

x10_state_median_income <- get_acs(geography = "state",
              variables = c(median_income = "B19013_001"),
              state = "MD",
              year = 2010)
```

Time to join the years and also mutate some percentage change formulas.

```
state_med_income <- x20_state_median_income %>%
  inner_join(x10_state_median_income, by = c("GEOID", "variable")) %>%
  rename(estimate_2020 = estimate.x,
         estimate_2010 = estimate.y,
         name = "NAME.x") %>%
```

```
    mutate(pct_change = (percent((estimate_2020-estimate_2010)/estimate_2010))) %>%
    select(GEOID, name, variable, estimate_2020, estimate_2010,pct_change)

  head(state_med_income)
```

```
# A tibble: 1 x 6
  GEOID name     variable      estimate_2020 estimate_2010 pct_change
  <chr> <chr>    <chr>                 <dbl>         <dbl> <formttbl>
1 24    Maryland median_income         87063         70647 23.24%
```

```
  #The household median income has increased by about 23%, which is a similar trend to that
```

I also wanted to find how the Key Bridge collapse could affect traffic patterns in Baltimore. While there isn't data yet about how the Key Bridge collapse has affected traffic patterns in the area. I found some data from the Maryland Department of Transportation about traffic patterns from the past few years, and will download it below.

```
  md_traffic <- read_csv("mdot_daily_traffic.csv") %>%
    clean_names()

  print(md_traffic)
```

```
# A tibble: 10,013 x 60
   gis_object_id station_id  county_code county_name  municipal_code
           <dbl> <chr>             <dbl> <chr>                 <dbl>
 1         84093 S2011020850           2 Anne Arundel              0
 2         84094 S2006010024           1 Allegany                  0
 3         84095 S2012030011           3 Baltimore                 0
 4         84096 S2012030219           3 Baltimore                 0
 5         84097 B3824                 5 Caroline                  0
 6         84098 B100070              10 Frederick                 0
 7         84099 S2011020816           2 Anne Arundel              0
 8         84100 B0645                 2 Anne Arundel              0
 9         84101 B030149               3 Baltimore                 0
10         84102 B0908                 3 Baltimore                 0
# i 10,003 more rows
# i 55 more variables: municipality_name <chr>, road_name <chr>,
#   route_prefix <chr>, route_number <dbl>, route_suffix <chr>,
#   milepoint <dbl>, begin_section <dbl>, end_section <dbl>,
#   station_description <chr>, road_section <chr>, rural_urban <chr>,
```

```
#   functional_class_code <dbl>, functional_class <chr>, route_id_legacy <chr>,
#   route_id <lgl>, mainline <dbl>, peak_hour_direction <dbl>, ...
```

In this dataframe, there are at least 60 columns. While this is a lot, I'm only going to be using a few of them. Here are some definitions of the columns based on this data definition notebook

Here's the column breakdown: gis_object_id is the location_id, or unique ID for the location of the traffic in an area. county_name is just the name of the county where this trip was recorded, or where it was taken. rural_urban just describes if the traffic run was in a rural or urban area. functional_class describes the kind of road or highway that the vehicle was driven on, such as an interstate or other kind of county road. all of the aadt_year values are the annual average daily traffic for any given day in that year. aadwt_year values are all of the annual average daily weekday traffic for that year. the current columns are the most up to date annual average daily weekday/daily traffic.

Now, let's clean this data of all of the columns we have.

```
#cleaning for names, and also grabbing all values that pertain to interstates, because the
md_traffic_clean <- md_traffic %>%
  select(gis_object_id, county_name, rural_urban, functional_class, aadt_2013, aadt_2014,
  filter(county_name == "Baltimore City")

view(md_traffic_clean)
```

I chose all of the traffic information with people going in and out of Baltimore County, where people are most likely to be impacted by the Key Bridge collapse. While Baltimore City and Baltimore County are two different areas, the traffic in the overall area will be affected regardless because of the interstates going in and out of the city. According to some sources, the roads to be most impacted by the Key Bridge Collapse are the interstates, including I-95, I-895 and MD-295. I want to explore how traffic has changed in these areas before the bridge collapse just to check on what could come with the collapse.

Now, let's make a percentage change and analyze the difference in years between daily and weekly traffic.

```
interstate_balt <- md_traffic_clean %>%
  mutate(aadt_pct_change = percent((aadt_current-aadt_2013)/aadt_2013)) %>%
  mutate(aawdt_pct_change = percent((aawdt_current-aawdt_2013)/aawdt_2013)) %>%
  select(gis_object_id, functional_class, county_name, aadt_2013, aadt_current, aawdt_2013
  filter(functional_class == "Interstate") %>%
  arrange(desc(aadt_pct_change))
```

```
view(interstate_balt)

#since 2013, multiple routes on the interstates have seen increased traffic by more than 1
```

Next, here's a visualization of the interstates around Baltimore City and where the Key Bridge collapsed around it. Because I-695 is heavily impacted by the Key Bridge collapse and traffic has increased just in the past 10 years, the other interstate routes are expected to pick up a lot more traffic.

```
cat('<iframe title="Interstates around Baltimore City could see more traffic after Key Bri
</script>')
```

Memo I, traffic Traffic on interstates in and around Baltimore City have increased by more than 200% on some routes since 2013, according to an analysis of Maryland Department of Transportation data. I-695, one of the interstates, currently has an access point cut off because of the Key Bridge collapse. While an about 200% traffic increase can be because of a multitude of factors such as business growth and higher density of people going in and out of the city, the recent collapse of the Key Bridge could make traffic on many other connecting interstates higher. The data, which shows average daily traffic and average weekday traffic in Maryland from 2013 to the current year, can be sorted by route and type of road. To check the growth of average daily and weekday traffic, I did a percent change formula and mapped out all the interstates surrounding Baltimore City. With this, I was able to find that multiple interstates could and would be impacted by the Key Bridge collapse, especially seeing previous rates of increased traffic. This could make roads more condensed and cause even more traffic around the city. While the data shows this increase, it is not able to say which specific interstates these traffic counts were completed. This is a limit that baseline reporting could instead fix by using different interstates and routes and checking with Department of Transportation officials, as well as commuters who use the interstates around the city. I would talk with community members from a multitude of communities around the city, especially those who work in maintenance and other industrial jobs. While many people in office jobs would be severely impacted by the increased amount of traffic, employees who work on homes, buildings and other elements of the city could have a harder time getting to their destinations to complete their jobs. I would find and speak with a lot of these employees, especially since they'll likely travel from other areas in the surrounding Baltimore areas. Especially those who live around Dundalk, Edgemere, Hart Miller Island and North Point Village, where residents will likely face increased impacts because of their proximity to the bridge.

Memo II, on Carroll County Support for Republican candidates has waned in Carroll County since 2012, according to an analysis of election data. Carroll County, which is one of 13 counties in Maryland that swings Republican, has instead seen a slight decrease in support for Republican candidates. In 2012, when nearly 65% favor for Republican candidates in 2012 to now about 60% in 2020 out of voters, the votes for Democratic candidates in Carroll County has increased. The data, scraped from the Maryland elections website, describes the

number of votes for each candidate, the percentage of voters based on county and the type of ballot used. While it doesn't entirely embody the demographics of those who voted for each candidate, it does describe the percentage of voters based on county. Carroll County, which is a majority white county, has also increased diversity since 2010. While it cannot be determined that support for Republican candidates decreased with the increased diversity in the county, it could be an interesting avenue to explore alongside other voting methods in the county. The data does not describe this based on census tract within the county, but with more on the ground reporting, one could piece together a narrative on how voting in the county has changed over time. I would reach out to election officials, senators, county council members and city council members to see how interest in voting and general election voting processes have changed in the county. I would also turn to other source of voting data such as federal election data to cross reference findings in Maryland data. I would also continue to calculate this information on my own to make sure the election data is accurate in its percentage calculations. I would also contact members of voting unions and other organizations to promote voting. There are multiple nonprofits that operate on a national level, but I want to check if they're reaching voters on the smaller, local level in counties. Is everyone in Carroll County knowledgeable about candidates, and how to vote? Are there any barriers to access for their voting? These are some baseline questions I'd like to answer, as well as in a writeup of how voting and elections in Carroll County have changed over time as we approach the 2024 general election.