NHIS Data 2020-2021 (Anxiety & Depression)

CJ Paine

2022-09-21

R Markdown

This is an R Markdown document following the steps I took to explore NHIS (National Health Interview Survey) data from 2020 and 2021, where I specifically look at data pertaining to the rate of diagnosed anxiety and depression among adults in the United States.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.0      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
adult2020 <- read_csv("adult20.csv")
```

```
## Warning: One or more parsing issues, see `problems()` for details
```

```
## Rows: 31568 Columns: 617
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (1): HHX
## dbl (605): URBRRL, RATCAT_A, INCGRP_A, INCTCFLG_A, FAMINCTC_A, IMPINCFLG_A, ...
## lgl  (11): OGFLG_A, OPFLG_A, CHFLG_A, PRPLCOV2_C_A, STOMAAGETC_A, RECTUAGETC...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
adult2021 <- read_csv("adult21.csv")
```

```
## Warning: One or more parsing issues, see `problems()` for details
```

```
## Rows: 29482 Columns: 622
## -- Column specification --------------------------------------------------------
## Delimiter: ","
```

```
## chr   (1): HHX
## dbl (599): URBRRL, RATCAT_A, IMPINCFLG_A, CVDVAC2YR_A, CVDVAC2MR_A, CVDVAC1Y...
## lgl  (22): OGFLG_A, OPFLG_A, CHFLG_A, MAFLG_A, PRPLCOV1_C_A, PRPLCOV2_C_A, P...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Cleaning the Data

```
index <- is.na(adult2020)
adult2020[index] <- 0
index2 <- is.na(adult2021)
adult2021[index2] <- 0

adult2020 <- adult2020 %>%
  select(ANXEV_A, DEPEV_A)
adult2021 <- adult2021 %>%
  select(ANXEV_A, DEPEV_A)
```

Tables for 2020 Data After cleaning my data and sorting columns, I create tables to explore the 2020 data.

This first table shows results for the question, "Have you EVER been told by a doctor or other health professional that you had any type of anxiety disorder?"

```
anxtable1<- adult2020 %>%
  count(ANXEV_A)
vec1 <- c('Yes', 'No', 'Refused', 'Dont Know')
anxtable1 <- cbind(anxtable1, vec1)
anxtable1 <- anxtable1 %>%
  mutate('Percent'=(n/31568)*100) %>%
  set_names(c('Code','Results','Description', 'Percent'))
print(anxtable1)
```

```
##   Code Results Description     Percent
## 1    1    4900         Yes 15.52204764
## 2    2   26617          No 84.31639635
## 3    7      24     Refused  0.07602636
## 4    9      27   Dont Know  0.08552965
```

This second table shows results for the question, "Have you EVER been told by a doctor or other health professional that you had any type of depression?"

```
deptable1 <- adult2020 %>%
  count(DEPEV_A)
deptable1 <- cbind(deptable1, vec1)
deptable1 <- deptable1 %>%
  mutate('Percent'=(n/31568)*100) %>%
  set_names(c('Code','Results','Description', 'Percent'))
print(deptable1)
```

```
##   Code Results Description     Percent
## 1    1    5518         Yes 17.47972631
## 2    2   25994          No 82.34287886
## 3    7      30     Refused  0.09503294
## 4    9      26   Dont Know  0.08236189
```

Tables for 2021 Data This first table shows results for the question, "Have you EVER been told by a doctor or other health professional that you had any type of anxiety disorder?"

```
anxtable2<- adult2021 %>%
  count(ANXEV_A)
vec1 <- c('Yes', 'No', 'Refused', 'Dont Know')
anxtable2 <- cbind(anxtable2, vec1)
anxtable2 <- anxtable2 %>%
  mutate('Percent'=(n/29482)*100) %>%
  set_names(c('Code','Results','Description', 'Percent'))
print(anxtable2)
```

```
##   Code Results Description    Percent
## 1    1    4860         Yes 16.4846347
## 2    2   24582          No 83.3796893
## 3    7      23     Refused  0.0780137
## 4    9      17   Dont Know  0.0576623
```

This second table shows results for the question, "Have you EVER been told by a doctor or other health professional that you had any type of depression?"

```
deptable2 <- adult2021 %>%
  count(DEPEV_A)
head(deptable2)
```

```
## # A tibble: 4 x 2
##   DEPEV_A     n
##     <dbl> <int>
## 1       1  5367
## 2       2 24063
## 3       7    23
## 4       9    29
```

```
deptable2 <- cbind(deptable2, vec1)
deptable2 <- deptable2 %>%
  mutate('Percent'=(n/29482)*100) %>%
  set_names(c('Code','Results','Description', 'Percent'))
print(deptable2)
```

```
##   Code Results Description    Percent
## 1    1    5367         Yes 18.2043281
## 2    2   24063          No 81.6192931
## 3    7      23     Refused  0.0780137
## 4    9      29   Dont Know  0.0983651
```

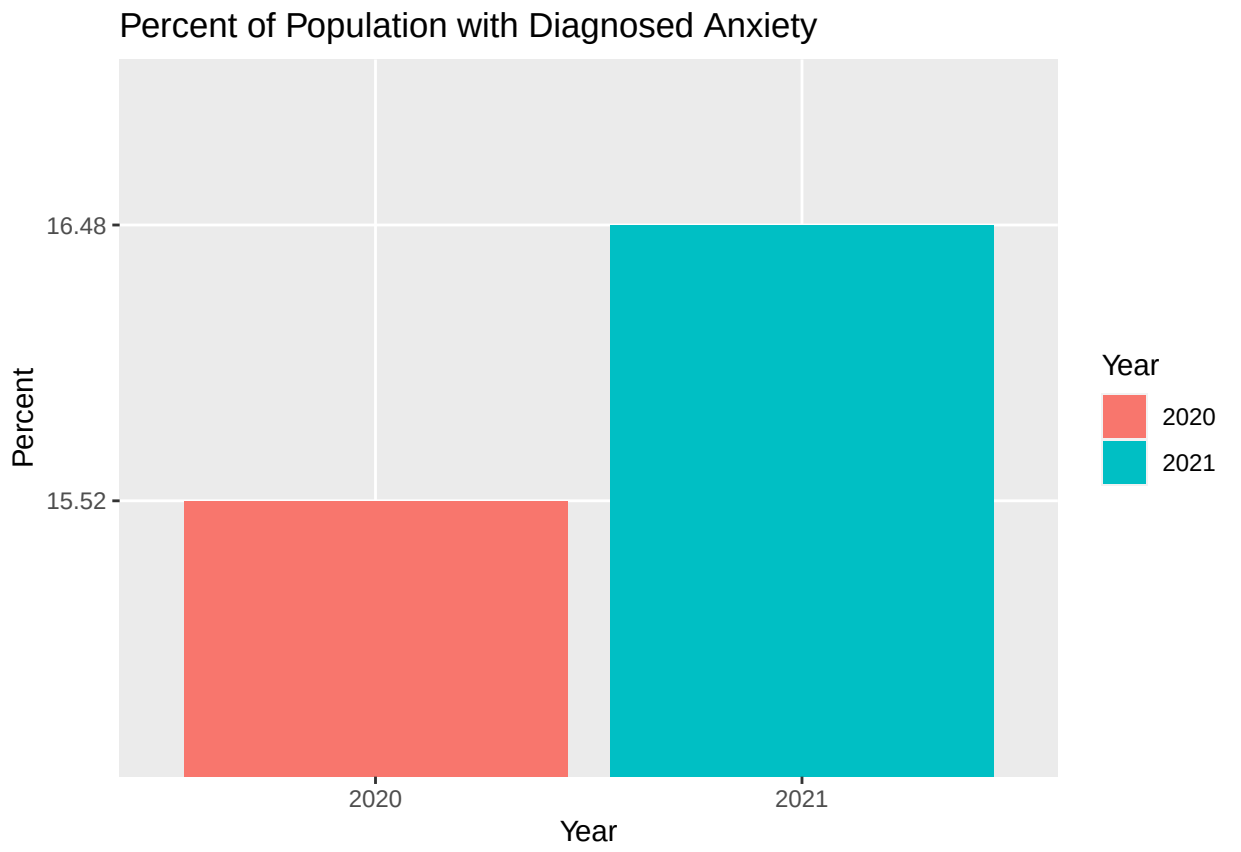
Creating a Data Frame for Visual Now that we have found the information we need to build a visual, let's first build the data frame for our visual.

```
vec3 = c('2020', '2021')
vec4 = c('15.52', '16.48')
vec5 = c('17.48', '18.20')
dataframe1 <- data.frame(vec3, vec4, vec5)
colnames(dataframe1)<-c('Year', 'PercAnx', 'PercDepr')
print(dataframe1)
```

```
##   Year PercAnx PercDepr
## 1 2020   15.52    17.48
## 2 2021   16.48    18.20
```

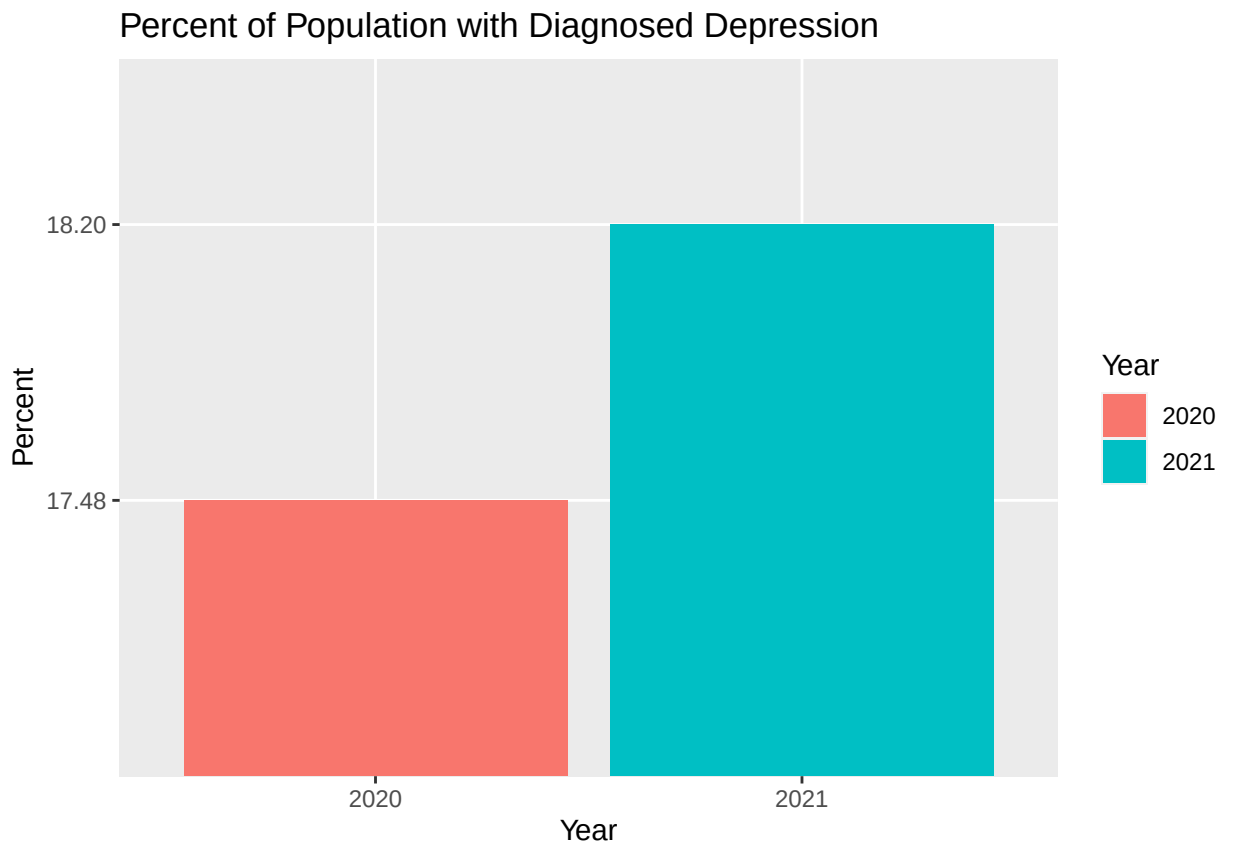
Visuals

```
dataframe1 %>%
  ggplot()+ geom_bar(stat='identity',mapping=aes(x=Year,fill=Year,y=PercAnx))+
  labs(title = "Percent of Population with Diagnosed Anxiety")+
  xlab("Year")+
  ylab("Percent")
```



Source: National Health Interview Survey

```
dataframe1 %>%
  ggplot()+ geom_bar(stat='identity',mapping=aes(x=Year,fill=Year,y=PercDepr))+
  labs(title = "Percent of Population with Diagnosed Depression")+
  xlab("Year")+
  ylab("Percent")
```



Source: National Health Interview Survey

In conclusion, it seems the percentage of the U.S. population with diagnosed anxiety and/or depression has grown from 2020 to 2021.