Using R to Replicate NCHS Data Brief No. 368

Prevalence of Tooth Loss Among Older Adults: United States, 2015–2018.

Data Brief No. 368

Load packages.

```
library(gt)
library(haven)
library(reactable)
library(srvyr)
library(survey)
library(tidyverse)
```

Read demographic files keeping variables of interest.

```
DEMO_I <- read_xpt("https://wwwn.cdc.gov/nchs/nhanes/2015-2016/demo_i.xpt") %>%
          select(SEQN, SDDSRVYR, RIAGENDR, RIDAGEYR, RIDRETH3, DMDEDUC2, SDMVSTRA, SDMVPSU, WTMEC2YR)
DEMO_J <- read_xpt("https://wwwn.cdc.gov/nchs/nhanes/2017-2018/demo_j.xpt") %>%
          select(SEQN, SDDSRVYR, RIAGENDR, RIDAGEYR, RIDRETH3, DMDEDUC2, SDMVSTRA, SDMVPSU, WTMEC2YR)
```

Append demographic files and create new variables. Make race/ethnicity and education factors to order the same as NCHS.

```
DEMO <- bind_rows(DEMO_I, DEMO_J) %>%
        mutate(
          gender = case_when(RIAGENDR == 1 ~ "Men",
                     RIAGENDR == 2 ~ "Women"),
          age = case_when(RIDAGEYR < 65 ~ "Less than 65",
                  RIDAGEYR < 70 ~ "65-69",
                  RIDAGEYR < 75 ~ "70-74",
                  TRUE ~ "75 and over"),
          race_ethnicity = case_when(RIDRETH3 < 3 ~ "Hispanic",
                             RIDRETH3 == 3 ~ "Non-Hispanic white",
                             RIDRETH3 == 4 ~ "Non-Hispanic black"),
          race_ethnicity = factor(race_ethnicity,
                             levels = c("Non-Hispanic white",
                               "Non-Hispanic black",
                               "Hispanic")),
          education = case_when(DMDEDUC2 %in% 1:2 ~ "Less than high school education",
                        DMDEDUC2 %in% 3:5 ~ "High school education or greater"),
          education = factor(education,
                        levels = c("Less than high school education",
                          "High school education or greater"))) %>%
        select(SEQN, gender, age, race_ethnicity, education, SDMVSTRA, SDMVPSU, WTMEC2YR)
```

Read oral health files keeping variables of interest.

```
OHXDEN_I <- read_xpt("https://wwwn.cdc.gov/nchs/nhanes/2015-2016/ohxden_i.xpt") %>%
            select(SEQN, OHX02TC:OHX31TC)
OHXDEN_J <- read_xpt("https://wwwn.cdc.gov/nchs/nhanes/2017-2018/ohxden_j.xpt") %>%
            select(SEQN, OHX02TC:OHX31TC)
```

Append oral health files and create variable for edentulism or complete tooth loss.

```
OHXDEN <- bind_rows(OHXDEN_I, OHXDEN_J) %>%
          mutate(
            tc = str_c(OHX02TC, OHX03TC, OHX04TC, OHX05TC, OHX06TC, OHX07TC, OHX08TC,
                   OHX09TC, OHX10TC, OHX11TC, OHX12TC, OHX13TC, OHX14TC, OHX15TC,
                   OHX18TC, OHX19TC, OHX20TC, OHX21TC, OHX22TC, OHX23TC, OHX24TC,
                   OHX25TC, OHX26TC, OHX27TC, OHX28TC, OHX29TC, OHX30TC, OHX31TC),
            edentulism = case_when(str_detect(tc, "[4]{28}") ~ 1,
                           str_detect(tc, "[1-5]{28}") ~ 0)) %>%
          select(SEQN, edentulism)
```

Join demographic and oral health data.

```
One <- left_join(DEMO, OHXDEN, by = "SEQN")
```

Define survey design.

```
NHANES <- One %>%
          as_survey_design(id = SDMVPSU, strata = SDMVSTRA, nest = TRUE, weight = WTMEC2YR)
```

Get data for Figure 1.

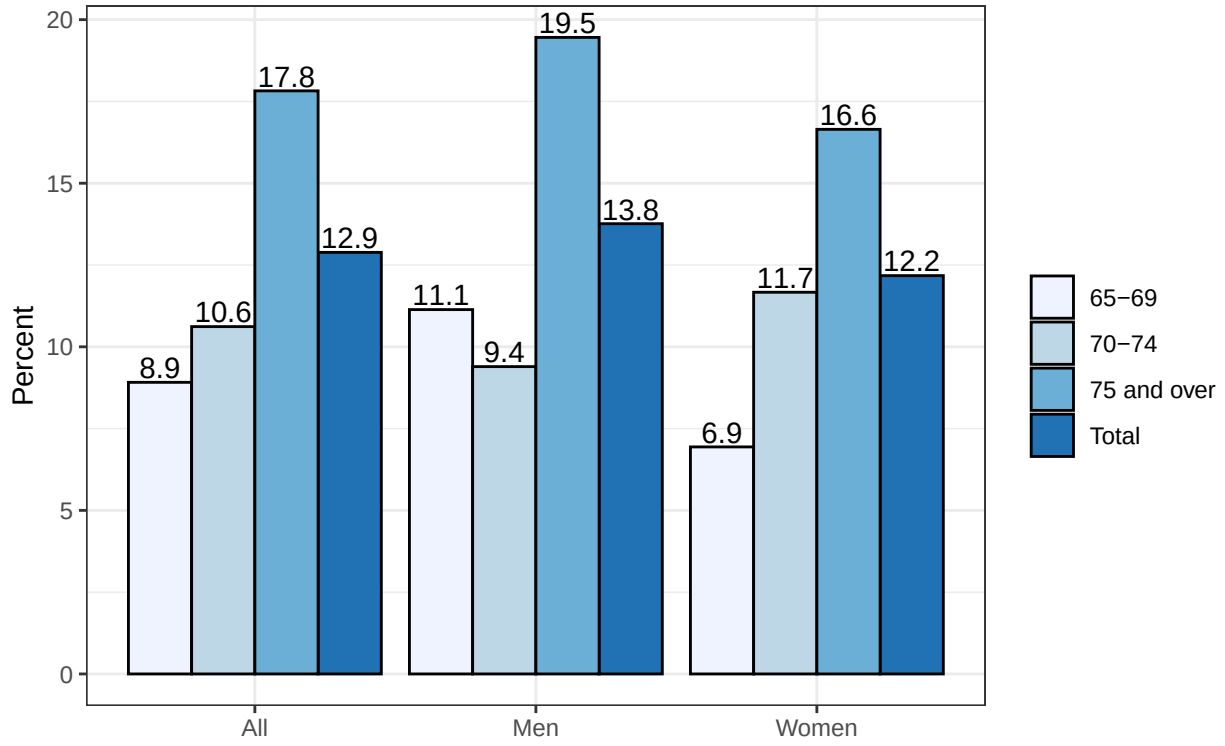
```
t1 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism)) %>%
      summarize(
        gender = "All",
        age = "Total",
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

t2 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism)) %>%
      group_by(age) %>%
      summarize(
        gender = "All",
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

t3 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism)) %>%
      group_by(gender) %>%
      summarize(
        age = "Total",
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

t4 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism)) %>%
```

```
      group_by(gender, age) %>%
      summarize(
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

table1 <- bind_rows(t1, t2, t3, t4)
```

Create Figure 1.

```
ggplot(table1, aes(gender, percent, fill = age)) +
  geom_bar(stat = "identity", position = position_dodge(), color = "black") +
  geom_text(aes(label = sprintf("%.1f", percent)), position = position_dodge(0.9), vjust = -0.2) +
  labs(title = paste("Figure 1. Prevalence of complete tooth loss among adults aged 65 and",
                     "over, by sex and age: United States, 2015-2018", sep = "\n"),
       x = element_blank(),
       y = "Percent",
       fill = element_blank()) +
  scale_fill_brewer(palette = "Blues") +
  theme_bw()
```

Figure 1. Prevalence of complete tooth loss among adults aged 65 and over, by sex and age: United States, 2015–2018



Create data table for Figure 1.

```
table1 %>%
  pivot_wider(names_from = gender, values_from = c(n, percent, percent_se)) %>%
  select(age, n_All, percent_All, percent_se_All, n_Men, percent_Men, percent_se_Men,
    n_Women, percent_Women, percent_se_Women) %>%
```

```
  gt() %>%
    tab_spanner(
      label = "All",
      columns = c(n_All, percent_All, percent_se_All)) %>%
    tab_spanner(
      label = "Men",
      columns = c(n_Men, percent_Men, percent_se_Men)) %>%
    tab_spanner(
      label = "Women",
      columns = c(n_Women, percent_Women, percent_se_Women)) %>%
    fmt_number(
      columns = c(n_All, n_Men, n_Women),
      decimals = 0) %>%
    fmt_number(
      columns = c(percent_All, percent_se_All, percent_Men, percent_se_Men, percent_Women, percent_se_W
      decimals = 1) %>%
    cols_label(
      age = "Age",
      n_All = "n", percent_All = "Percent", percent_se_All = "SE",
      n_Men = "n", percent_Men = "Percent", percent_se_Men = "SE",
      n_Women = "n", percent_Women = "Percent", percent_se_Women = "SE")
```

	All			Men			Women		
Age	n	Percent	SE	n	Percent	SE	n	Percent	SE
Total	2,583	12.9	1.3	1,301	13.8	1.6	1,282	12.2	1.4
65-69	808	8.9	1.6	399	11.1	2.4	409	6.9	1.7
70-74	637	10.6	1.5	339	9.4	1.7	298	11.7	2.1
75 and over	1,138	17.8	1.5	563	19.5	1.8	575	16.6	2.0

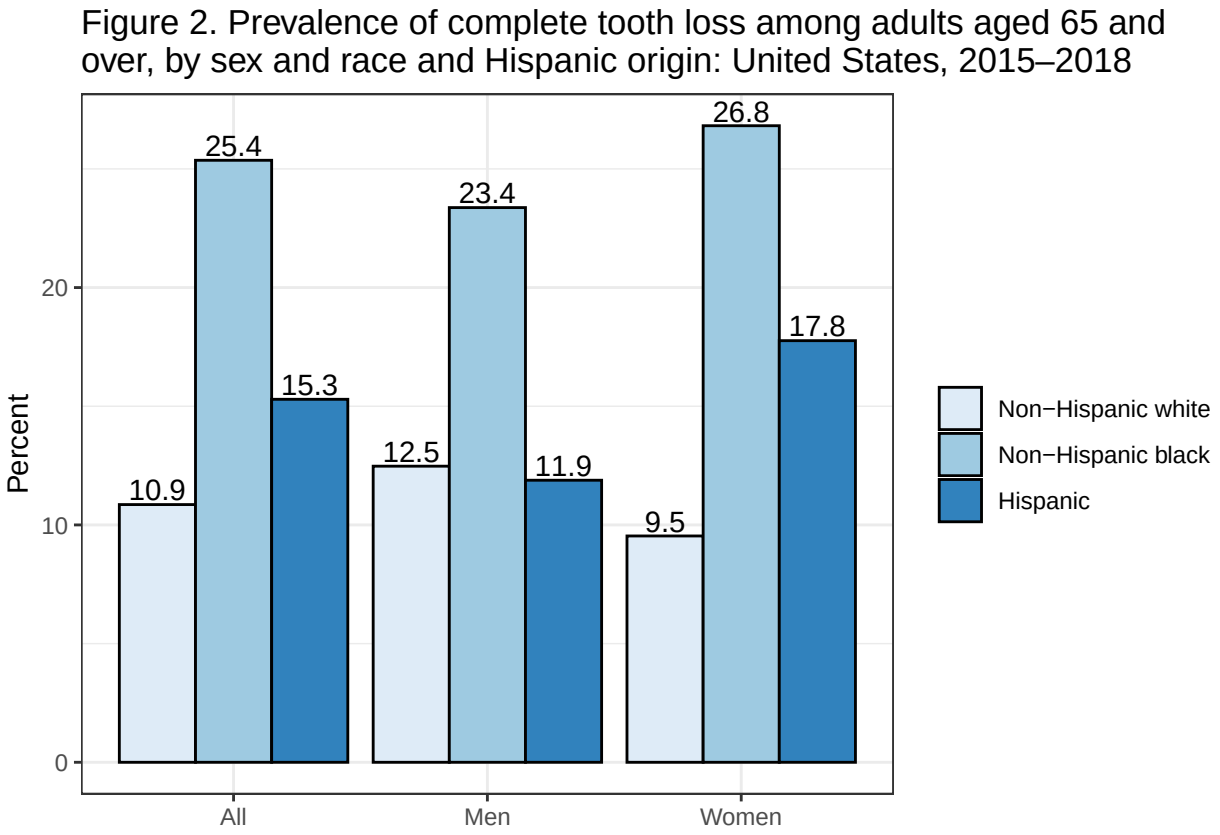
Get data for Figure 2.

```
t1 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism), !is.na(race_ethnicity)) %>%
      group_by(race_ethnicity) %>%
      summarize(
        gender = "All",
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

t2 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism), !is.na(race_ethnicity)) %>%
      group_by(gender, race_ethnicity) %>%
      summarize(
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

table2 <- bind_rows(t1, t2)
```

Create Figure 2.

```
ggplot(table2, aes(gender, percent, fill = race_ethnicity)) +
  geom_bar(stat = "identity", position = position_dodge(), color = "black") +
```

```
  geom_text(aes(label = sprintf("%.1f", percent)), position = position_dodge(0.9), vjust = -0.2) +
  labs(title = paste("Figure 2. Prevalence of complete tooth loss among adults aged 65 and",
                     "over, by sex and race and Hispanic origin: United States, 2015-2018", sep = "\n"),
       x = element_blank(),
       y = "Percent",
       fill = element_blank()) +
  scale_fill_brewer(palette = "Blues") +
  theme_bw()
```



Create data table for Figure 2.

```
reactable(table2 %>%
  pivot_wider(names_from = gender, values_from = c(n, percent, percent_se)) %>%
    select(race_ethnicity, n_All, percent_All, percent_se_All, n_Men, percent_Men, percent_se_Men,
      n_Women, percent_Women, percent_se_Women),
  columns = list(
    race_ethnicity = colDef(name = "Race & Hispanic Origin", minWidth = 200),
    n_All = colDef(name = "n", format = colFormat(separators = TRUE)),
    percent_All = colDef(name = "Percent", format = colFormat(digits = 1)),
    percent_se_All = colDef(name = "SE", format = colFormat(digits = 1)),
    n_Men = colDef(name = "n", format = colFormat(separators = TRUE)),
    percent_Men = colDef(name = "Percent", format = colFormat(digits = 1)),
    percent_se_Men = colDef(name = "SE", format = colFormat(digits = 1)),
    n_Women = colDef(name = "n", format = colFormat(separators = TRUE)),
    percent_Women = colDef(name = "Percent", format = colFormat(digits = 1)),
    percent_se_Women = colDef(name = "SE", format = colFormat(digits = 1))),
```

```
  columnGroups = list(
    colGroup(name = "All", columns = c("n_All", "percent_All", "percent_se_All")),
    colGroup(name = "Men", columns = c("n_Men", "percent_Men", "percent_se_Men")),
    colGroup(name = "Women", columns = c("n_Women", "percent_Women", "percent_se_Women"))),
  defaultColDef = colDef(minWidth = 75),
  striped = TRUE)
```

	All			Men			
Race & Hispanic Origin	**n**	**Percent**	**SE**	**n**	**Percent**	**SE**	
Non-Hispanic white	1217	10.8571 5226782 29	1.59972 7443007 42	641	12.4756 1297776 98	1.91265 2300095 34	5
Non-Hispanic black	510	25.3628 0950685 72	2.75272 4092659 18	263	23.3696 4102962 52	2.79447 5382674 73	2
Hispanic	554	15.2932 6601000 97	1.84051 9164000 8	249	11.88320 7046731 3	2.481132 6632163 9	3

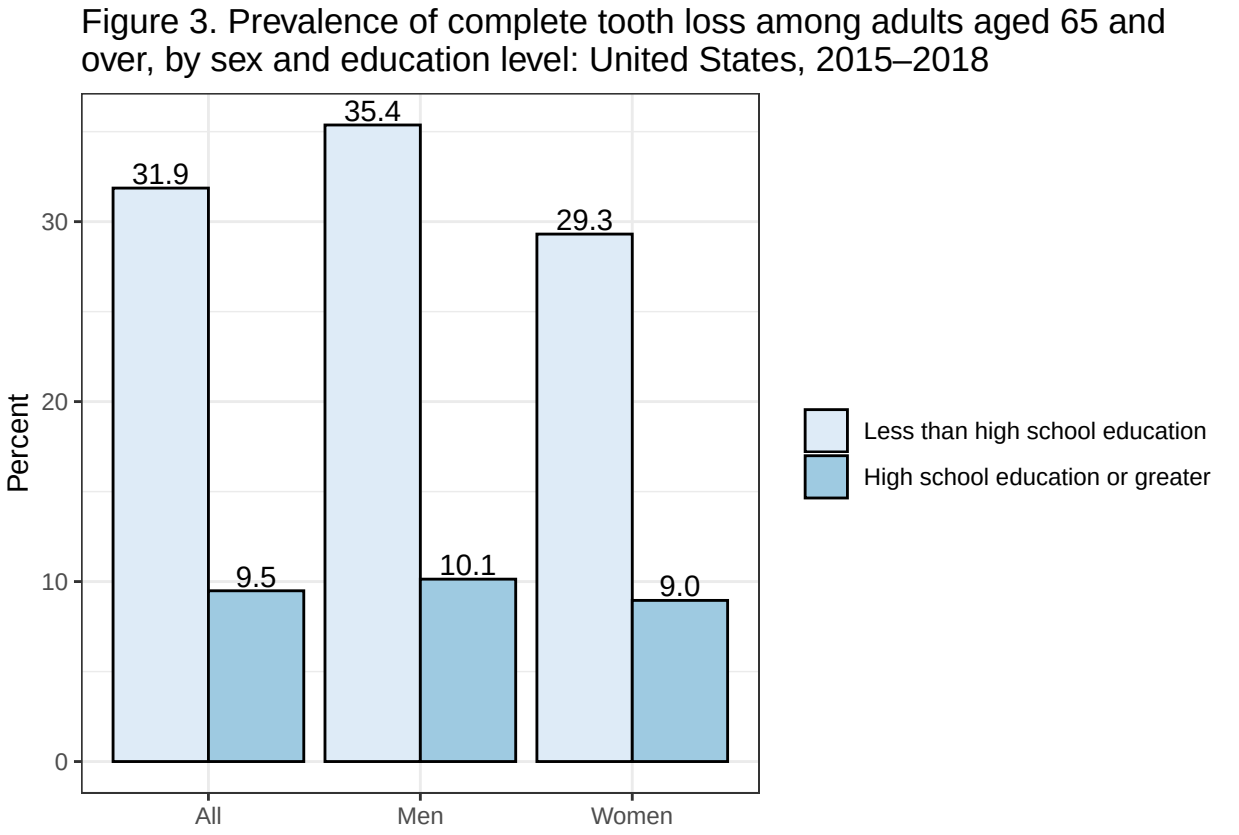
Get data for Figure 3.

```
t1 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism), !is.na(education)) %>%
      group_by(education) %>%
      summarize(
        gender = "All",
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)

t2 <- NHANES %>% filter(age != "Less than 65", !is.na(edentulism), !is.na(education)) %>%
      group_by(gender, education) %>%
      summarize(
        n = unweighted(n()),
        percent = survey_mean(edentulism) * 100)
```

```
table3 <- bind_rows(t1, t2)
```

Create Figure 3.

```
ggplot(table3, aes(gender, percent, fill = education)) +
  geom_bar(stat = "identity", position = position_dodge(), color = "black") +
  geom_text(aes(label = sprintf("%.1f", percent)), position = position_dodge(0.9), vjust = -0.2) +
  labs(title = paste("Figure 3. Prevalence of complete tooth loss among adults aged 65 and",
                     "over, by sex and education level: United States, 2015-2018", sep = "\n"),
       x = element_blank(),
       y = "Percent",
       fill = element_blank()) +
  scale_fill_brewer(palette = "Blues") +
  theme_bw()
```



Create data table for Figure 3.
(Note: The standard error for women with less than a high school education is 3.7 in the Data Brief. Assume difference due to rounding or typo.)

```
reactable(table3 %>%
  pivot_wider(names_from = gender, values_from = c(n, percent, percent_se)) %>%
    select(education, n_All, percent_All, percent_se_All, n_Men, percent_Men, percent_se_Men,
      n_Women, percent_Women, percent_se_Women),
  columns = list(
    education = colDef(name = "Education Level", minWidth = 225),
    n_All = colDef(name = "n", format = colFormat(separators = TRUE)),
```

```
  percent_All = colDef(name = "Percent", format = colFormat(digits = 1)),
  percent_se_All = colDef(name = "SE", format = colFormat(digits = 1)),
  n_Men = colDef(name = "n", format = colFormat(separators = TRUE)),
  percent_Men = colDef(name = "Percent", format = colFormat(digits = 1)),
  percent_se_Men = colDef(name = "SE", format = colFormat(digits = 1)),
  n_Women = colDef(name = "n", format = colFormat(separators = TRUE)),
  percent_Women = colDef(name = "Percent", format = colFormat(digits = 1)),
  percent_se_Women = colDef(name = "SE", format = colFormat(digits = 1))),
columnGroups = list(
  colGroup(name = "All", columns = c("n_All", "percent_All", "percent_se_All")),
  colGroup(name = "Men", columns = c("n_Men", "percent_Men", "percent_se_Men")),
  colGroup(name = "Women", columns = c("n_Women", "percent_Women", "percent_se_Women"))),
defaultColDef = colDef(minWidth = 75),
striped = TRUE)
```

	All			Men		
Education Level	**n**	**Percent**	**SE**	**n**	**Percent**	**SE**
Less than high school education	696	31.86506138056045	3.1311240365685 3	335	35.36992957566 07	4.041958707138 58
High school education or greater	1876	9.488499047557 42	1.208992399384 57	960	10.13557566604 91	1.478063364424 32