> GSS2006 <- read\_csv("GSS\_2006.csv")

Parsed with column specification:

cols(

.default = col\_character(),

prestg10 = col\_integer(),

prestg105plus = col\_integer(),

sppres10 = col\_integer(),

sppres105plus = col\_integer(),

papres10 = col\_integer(),

papres105plus = col\_integer(),

mapres10 = col\_integer(),

mapres105plus = col\_integer(),

sei10 = col\_double(),

spsei10 = col\_double(),

pasei10 = col\_double(),

masei10 = col\_double(),

sei10educ = col\_double(),

spsei10educ = col\_double(),

pasei10educ = col\_double(),

masei10educ = col\_double(),

sei10inc = col\_double(),

spsei10inc = col\_double(),

pasei10inc = col\_double(),

masei10inc = col\_double()

# ... with 110 more columns

)

See spec(...) for full column specifications.

Warning: 7 parsing failures.

row # A tibble: 5 x 5 col row col expected actual file expected <int> <chr> <chr> <chr> <chr> actual 1 1722 physhlth an integer DONT KNOW 'GSS\_2006.csv' file 2 2958 prozfor1 an integer DONT KNOW 'GSS\_2006.csv' row 3 3164 adults no trailing characters " or more" 'GSS\_2006.csv' col 4 3170 sphrs2 no trailing characters + hrs 'GSS\_2006.csv' expected 5 3246 physhlth an integer DONT KNOW 'GSS\_2006.csv'

... ................................. ... ................................................................. ........ ....................................... [... truncated]

Warning message:

In rbind(names(probs), probs\_f) :

number of columns of result is not a multiple of vector length (arg 1)

> # Select adults from our data frame and drop all null values

> gssadults\_06 <- GSS2006 %>%

+ select(adults, sex) %>%

+ filter(!is.na(adults)) %>%

* filter(!is.na(sex))

> # 80% Confidence Intervel

> gssadults\_06 %>%

+ summarise(mean.ad = mean(adults),

+ sd.ad = sd(adults),

+ n.ad = n()) %>%

+ mutate(se.ad = sd.ad / sqrt(n.ad),

+ "80%\_low.CI" = mean.ad - qt(1 - (0.20 / 2),

+ n.ad -1) \*se.ad,

+ "80%\_upp.CI" = mean.ad + qt(1 - (0.20 / 2),

+ n.ad -1) \*se.ad)

# A tibble: 1 x 6

mean.ad sd.ad n.ad se.ad `80%\_low.CI` `80%\_upp.CI`

<dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 1.86 0.793 4509 0.0118 1.84 1.87

> # 95% Confidence Intervel

> gssadults\_06 %>%

+ summarise(mean.ad = mean(adults),

+ sd.ad = sd(adults),

+ n.ad = n()) %>%

+ mutate(se.ad = sd.ad / sqrt(n.ad),

+ "95%\_low.CI" = mean.ad - qt(1 - (0.05 / 2), n.ad -1) \*se.ad,

+ "95%\_upp.CI" = mean.ad + qt(1 - (0.05 / 2), n.ad -1) \*se.ad)

# A tibble: 1 x 6

mean.ad sd.ad n.ad se.ad `95%\_low.CI` `95%\_upp.CI`

<dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 1.86 0.793 4509 0.0118 1.84 1.88

> # 99% Confidence Intervel

> gssadults\_06 %>%

+ summarise(mean.ad = mean(adults),

+ sd.ad = sd(adults),

+ n.ad = n()) %>%

+ mutate(se.ad = sd.ad / sqrt(n.ad),

+ "99%\_low.CI" = mean.ad - qt(1 - (0.01 / 2), n.ad -1) \*se.ad,

+ "99%\_upp.CI" = mean.ad + qt(1 - (0.01 / 2), n.ad -1) \*se.ad)

# A tibble: 1 x 6

mean.ad sd.ad n.ad se.ad `99%\_low.CI` `99%\_upp.CI`

<dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 1.86 0.793 4509 0.0118 1.83 1.89

> # 80% Confidence Intervel by `sex'

> gssadults\_06 %>%

+ group\_by(sex) %>% # this output gives us outputs based on `sex'

+ summarise(mean.ad = mean(adults),

+ sd.ad = sd(adults),

+ n.ad = n()) %>%

+ mutate(se.ad = sd.ad / sqrt(n.ad),

+ "80%\_low.CI" = mean.ad - qt(1 - (0.20 / 2),

+ n.ad -1) \*se.ad,

+ "80%\_upp.CI" = mean.ad + qt(1 - (0.20 / 2),

+ n.ad -1) \*se.ad)

# A tibble: 2 x 7

sex mean.ad sd.ad n.ad se.ad `80%\_low.CI` `80%\_upp.CI`

<chr> <dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 female 1.82 0.780 2507 0.0156 1.80 1.84

2 male 1.91 0.805 2002 0.0180 1.89 1.94

> # 95% Confidence Intervel by `sex'

> gssadults\_06 %>%

+ group\_by(sex) %>%

+ summarise(mean.ad = mean(adults),

+ sd.ad = sd(adults),

+ n.ad = n()) %>%

+ mutate(se.ad = sd.ad / sqrt(n.ad),

+ "95%\_low.CI" = mean.ad - qt(1 - (0.05 / 2), n.ad -1) \*se.ad,

+ "95%\_upp.CI" = mean.ad + qt(1 - (0.05 / 2), n.ad -1) \*se.ad)

# A tibble: 2 x 7

sex mean.ad sd.ad n.ad se.ad `95%\_low.CI` `95%\_upp.CI`

<chr> <dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 female 1.82 0.780 2507 0.0156 1.79 1.85

2 male 1.91 0.805 2002 0.0180 1.88 1.95

> # 99% Confidence Intervel by `sex'

> gssadults\_06 %>%

+ group\_by(sex) %>%

+ summarise(mean.ad = mean(adults),

+ sd.ad = sd(adults),

+ n.ad = n()) %>%

+ mutate(se.ad = sd.ad / sqrt(n.ad),

+ "99%\_low.CI" = mean.ad - qt(1 - (0.01 / 2), n.ad -1) \*se.ad,

+ "99%\_upp.CI" = mean.ad + qt(1 - (0.01 / 2), n.ad -1) \*se.ad)

# A tibble: 2 x 7

sex mean.ad sd.ad n.ad se.ad `99%\_low.CI` `99%\_upp.CI`

<chr> <dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 female 1.82 0.780 2507 0.0156 1.78 1.86

2 male 1.91 0.805 2002 0.0180 1.87 1.96