Problem Set 3 Answer Key

Joe Ornstein

2022-09-07

```
## Load Packages and Data --------------

library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.7     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(here)
```

```
## here() starts at C:/Users/jo22058/Documents/teaching/intro-political-methodology
```

```
ces_raw <- read_csv(here('data/raw/ces-2020/CES20_Common_OUTPUT_vv.csv'))
```

```
## New names:
## * `` -> `...1`
## Rows: 61000 Columns: 717
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (175): race_other, CC20_hisp_t, CC20_asian_t, CC20_309d_t, CC20_364a_t,...
## dbl  (537): ...1, caseid, commonweight, commonpostweight, vvweight, vvweight...
## lgl    (1): multrace_97
## dttm   (4): starttime, endtime, starttime_post, endtime_post
##
## 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.
```

```
## Problem 1 ----------
# Are labor union members more likely to be Democrats or Republicans?

ces_clean <- ces_raw |>
  mutate(union = case_when(union == 1 ~ 'Yes',
                           union == 2 ~ 'Former',
                           union == 3 ~ 'No')) |>
  mutate(partyid = case_when(pid3 == 1 ~ 'Democrat',
                             pid3 == 2 ~ 'Republican',
                             TRUE ~ 'Other'))
ces_clean |>
```

```
  group_by(union) |>
  summarize(pct_democrat = sum(partyid == 'Democrat') / n() * 100,
            count = n())
```

```
## # A tibble: 4 x 3
##   union  pct_democrat count
##   <chr>         <dbl> <int>
## 1 Former         39.4 10726
## 2 No             36.0 46370
## 3 Yes            47.0  3832
## 4 <NA>           29.2    72
```

```
## Problem 2 ----------------
# What is the median age of people with landline phones,
# compared to people who only have cell phones?

ces_clean <- ces_clean |>
  mutate(has_landline = case_when(phone == 1 ~ 'Has Landline',
                                  phone == 2 ~ 'Cell Only',
                                  phone == 3 ~ 'Has Landline',
                                  phone == 4 ~ 'Neither')) |>
  mutate(age = 2020 - birthyr)

ces_clean |>
  filter(!is.na(has_landline)) |>
  group_by(has_landline) |>
  summarize(median_age = median(age, na.rm = TRUE))
```

```
## # A tibble: 3 x 2
##   has_landline median_age
##   <chr>             <dbl>
## 1 Cell Only            42
## 2 Has Landline         60
## 3 Neither              29
```

```
## Problem 3 -------------------------
# Are the people who read newspapers more likely to correctly answer
# the questions about who controls the US House and Senate?

ces_clean |>
  count(CC20_300_3)
```

```
## # A tibble: 2 x 2
##   CC20_300_3     n
##        <dbl> <int>
## 1          1 22983
## 2          2 38017
```

```
ces_clean <- ces_clean |>
  mutate(reads_newspaper = (CC20_300_3 == 1)) |>
  mutate(
    answered_correctly_both =
      if_else(CC20_310a == 2 & CC20_310b == 1, 1, 0))

ces_clean |>
  filter(!is.na(reads_newspaper),
```

```
        !is.na(answered_correctly_both)) |>
  group_by(reads_newspaper) |>
  summarize(pct_correct = mean(answered_correctly_both) * 100)
```

```
## # A tibble: 2 x 2
##   reads_newspaper pct_correct
##   <lgl>                 <dbl>
## 1 FALSE                  50.4
## 2 TRUE                   75.5
```

```
## Problem 4 --------------------------------
# Which religious groups are the most and least likely to support
# making abortion illegal in all circumstances?

ces_clean <- ces_clean |>
  mutate(religion = case_when(religpew == 1 ~ 'Protestant',
                              religpew == 2 ~ 'Roman Catholic',
                              religpew == 3 ~ 'Mormon',
                              religpew == 4 ~ 'Orthodox',
                              religpew == 5 ~ 'Jewish',
                              religpew == 6 ~ 'Muslim',
                              religpew == 7 ~ 'Buddhist',
                              religpew == 8 ~ 'Hindu',
                              religpew == 9 ~ 'Atheist',
                              religpew == 10 ~ 'Agnostic',
                              religpew == 11 ~ 'Nothing in particular',
                              religpew == 12 ~ 'Something else')) |>
  mutate(abortion_illegal = if_else(CC20_332f == 1, 1, 0))

ces_clean |>
  filter(!is.na(religion)) |>
  group_by(religion) |>
  summarize(pct_support_abortion_ban = mean(abortion_illegal,
                                            na.rm = TRUE) * 100,
            num_respondents = n()) |>
  arrange(-pct_support_abortion_ban)
```

```
## # A tibble: 12 x 3
##    religion              pct_support_abortion_ban num_respondents
##    <chr>                                    <dbl>           <int>
##  1 Muslim                                   24.7              438
##  2 Protestant                               22.7            19295
##  3 Orthodox                                 21.1              352
##  4 Hindu                                    19.9              221
##  5 Something else                           18.8             4150
##  6 Roman Catholic                           18.4            11232
##  7 Mormon                                   16.9              763
##  8 Nothing in particular                    12.0            13682
##  9 Buddhist                                  9.43             583
## 10 Jewish                                    7.88            1560
## 11 Atheist                                   4.16            4642
## 12 Agnostic                                  3.58            4048
```