Group Project

Evan McKee, Roman, etc.

12/06/2022

Input data

```
## setting working directory
setwd("C:/git/statsGroupProject/R")
## input data
load("ces.RData")

library(ggplot2)
```

Variables to factor

```
ces$polengage <- ces$voted + ces$meeting + ces$sign +
    ces$campaign + ces$protest + ces$contact + ces$donate


## Region
ces$region <- factor(ces$region,
    levels = 1:4,
    labels = c("Northeast", "Midwest", "South", "West")
)

## Gender
ces$gender <- factor(ces$gender,
    levels = 1:2,
    labels = c("Male", "Female")
)

## Education
ces$educ <- factor(ces$educ,
    levels = 1:6,
    labels = c(
        "No HS", "High School Graduate", "Some college",
        "2-year degree", "4-year degree", "Post-grad"
    )
)

## Race
ces$race <- factor(ces$race,
```

```
    levels = 1:8,
    labels = c(
        "White", "Black", "Hispanic", "Asian",
        "Native American", "Middle Eastern", "Two or more races", "Other"
    )
)

## Hispanic
ces$hispanic <- factor(ces$hispanic,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Homeowner
ces$ownhome <- factor(ces$ownhome,
    levels = 1:3,
    labels = c("Own", "Rent", "Other")
)

## Neighborhood Type
ces$urbancity <- factor(ces$urbancity,
    levels = 1:5,
    labels = c("City", "Suburb", "Town", "Rural Area", "Other")
)

## Union
ces$unionhh <- factor(ces$unionhh,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Religious
ces$religious <- factor(ces$religious,
    levels = 1:4,
    labels = c(
        "Very important", "Not too important",
        "Not too important", "Not at all important"
    )
)

## Family income
ces$faminc <- factor(ces$faminc,
    levels = 1:16,
    labels = c(
        "< 10,000", "10,000 - 19,999", "20,000 - 29,999",
        "30,000 - 39,999", "40,000 - 49,999",
        "50,000 - 59,999", "60,000 - 69,999", "70,000 - 79,999",
        "80,000 - 99,999", "100,000-119,000", "120,000 - 149,999",
        "150,000 - 199,999", "200,000 - 249,999", "250,000 - 349,999",
        "350,000 - 499,999", "500,000 or more"
    )
)
```

```r
## Voted
ces$voted <- factor(ces$voted,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Meeting
ces$meeting <- factor(ces$meeting,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Signed
ces$sign <- factor(ces$sign,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Campaign
ces$campaign <- factor(ces$campaign,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Protest
ces$protest <- factor(ces$protest,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Contact
ces$contact <- factor(ces$contact,
    levels = 1:2,
    labels = c("Yes", "No")
)

## Donate
ces$donate <- factor(ces$donate,
    levels = 1:2,
    labels = c("Yes", "No")
)

## ideo5
ces$ideo5 <- factor(ces$ideo5,
    levels = 1:5,
    labels = c(
        "Very liberal", "Liberal", "Moderate",
        "Conservative", "Very Conservative"
    )
)

## pid3
ces$pid3 <- factor(ces$pid3,
```

```
    levels = 1:5,
    labels = c(
        "Democrat", "Republican", "Independent",
        "Other", "Not Sure"
    )
)

## pid7
ces$pid7 <- factor(ces$pid7,
    levels = 1:7,
    labels = c(
        "Strong Democrat", "Not very strong Democrat",
        "Lean Democract", "Independent", "Lean Republican",
        "Not very strong Republican", "Strong Republican"
    )
)

## medicare
ces$medicare <- factor(ces$medicare,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## ACA
ces$ACA <- factor(ces$ACA,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## abortion
ces$abortion <- factor(ces$abortion,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## EPA
ces$EPA <- factor(ces$EPA,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## Increase the number of police
ces$police_incr <- factor(ces$police_incr,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## Decrease the number of police
ces$police_decr <- factor(ces$police_decr,
    levels = 1:2,
    labels = c("Support", "Oppose")
)
```

```r
## Trade Tariffs
ces$trade <- factor(ces$trade,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## Raise minimum wage
ces$minwage <- factor(ces$minwage,
    levels = 1:2,
    labels = c("Support", "Oppose")
)

## Work requirement for food stamps
ces$work_req <- factor(ces$work_req,
    levels = 1:2,
    labels = c("Support", "Oppose")
)
```

```r
ces[1:2, ]
```

```
## # A tibble: 2 x 34
##    ...1 gender educ     race  hispa~1 region medic~2 ACA   abort~3 EPA   polic~4
##   <dbl> <fct>  <fct>    <fct> <fct>   <fct>  <fct>   <fct> <fct>   <fct> <fct>
## 1     1 Male   2-year ~ White No      North~ <NA>    <NA>  <NA>    Supp~ Support
## 2     2 Female Post-gr~ White No      South  Support <NA>  <NA>    Supp~ <NA>
## # ... with 23 more variables: police_decr <fct>, trade <fct>, minwage <fct>,
## #   work_req <fct>, residency <dbl>, pid3 <fct>, pid7 <fct>, ownhome <fct>,
## #   urbancity <fct>, unionhh <fct>, religious <fct>, ideo5 <fct>, faminc <fct>,
## #   voted <fct>, meeting <fct>, sign <fct>, campaign <fct>, protest <fct>,
## #   contact <fct>, donate <fct>, commonweight <dbl>, age <dbl>,
## #   polengage <dbl>, and abbreviated variable names 1: hispanic, 2: medicare,
## #   3: abortion, 4: police_incr
```

Histogram

```r
ggplot(ces, aes(polengage)) +
    geom_histogram(color = "#a8a8a8", fill = "#000000", binwidth = 1) +
    labs(
        title = "Distribution of Political Engagment",
        ## caption = "Source: Gapminder dataset",
        x = "Political Activities",
        y = "Count"
    ) +
    theme_classic() +
    theme(
        plot.title = element_text(color = "#0099F8", size = 16, face = "bold"),
```

```
        plot.subtitle = element_text(size = 10, face = "bold"),
        plot.caption = element_text(face = "italic")
    )
```

```
## Warning: Removed 9449 rows containing non-finite values (stat_bin).
```