Data Science for Public Policy

Assignment 04: Data Visualization, Quarto, and Git

Ronin McCobb and Daniela Nagar

Data Source: IPUMS

```r
1  library(tidyverse)
2  library(ipumsr)
3  library(srvyr)
4  library(haven)
```

```r
1  ddi <- read_ipums_ddi("data/cps_00004.xml")
2  data <- read_ipums_micro(ddi)
```

Graph 1: Food Security by State

```
1  #Graph 1
2
3  data |>
4    filter(FSRAWSCRA < 11 & FSRAWSCRA > 0 ) |>
5    ggplot(aes(x=FSRAWSCRA)) +
6    geom_bar(width = 1, fill = "blue") +
7     labs(
8        title = "Food Security Scores in the United States",
9        caption = "Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles,
10 J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards,
11 Megan Schouweiler, and Michael Westberry. IPUMS CPS Version 12.0 [dataset].
12 Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0"
13       ) +
14    xlab("Score") +
15    ylab("Count")
```
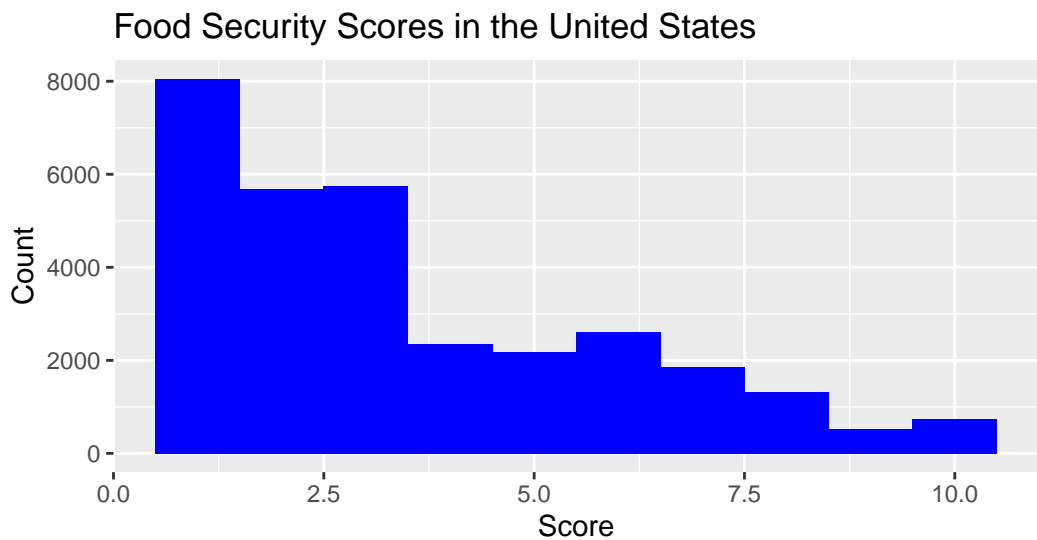
### Food Security Scores in the United States

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles,
J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards,
Megan Schouweiler, and Michael Westberry. IPUMS CPS Version 12.0 [dataset].
Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0

To start, we looked at the number of respondents across all states who have every food security score. Food security scores span from 1-10. Respondents with larger scores are more food insecure. The graph reflects what we would expect by trending downward as scores grow

higher, indicating that fewer people are lacking in all food security measures but many are lacking in at least one measure.

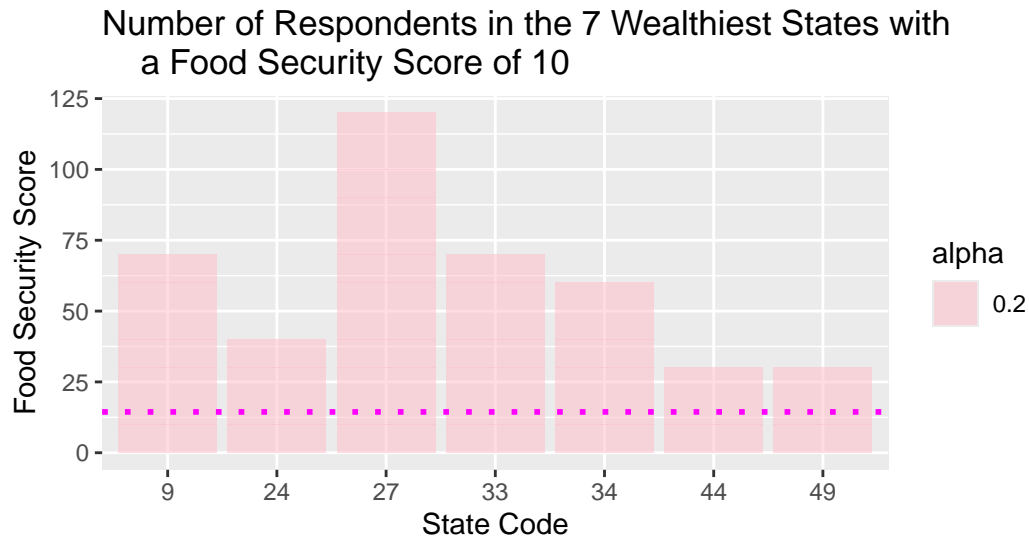## Graph 2: Income by State, finding the wealthiest

```
1   #Data Visualization 2: Average Family Income by State Among Wealthiest States
2   avg_faminc <- data |>
3     group_by(STATEFIP) |>
4     summarize(avg_faminc = mean(FAMINC, na.rm = TRUE)) |>
5     filter(avg_faminc >=785) |>
6   mutate(
7       STATEFIP = as.numeric(STATEFIP))
8
9   ggplot(avg_faminc, aes(x = STATEFIP, y = avg_faminc)) +
10    geom_point(color = "blue", size = 3, shape = 18) +
11    geom_text(aes(label = STATEFIP), vjust = -0.5, size = 3) +
12    labs (
13      title = "Average Family Income by State Among Wealthiest States",
14      x = "State Code",
15      y = "Average Family Income",
16       caption = "Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles,
17  J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards,
18  Megan Schouweiler, and Michael Westberry. IPUMS CPS Version 12.0 [dataset].
19  Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0"
20      )
```

Average Family Income by State Among Wealthiest States

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles,
J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards,
Megan Schouweiler, and Michael Westberry. IPUMS CPS Version 12.0 [dataset].
Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0
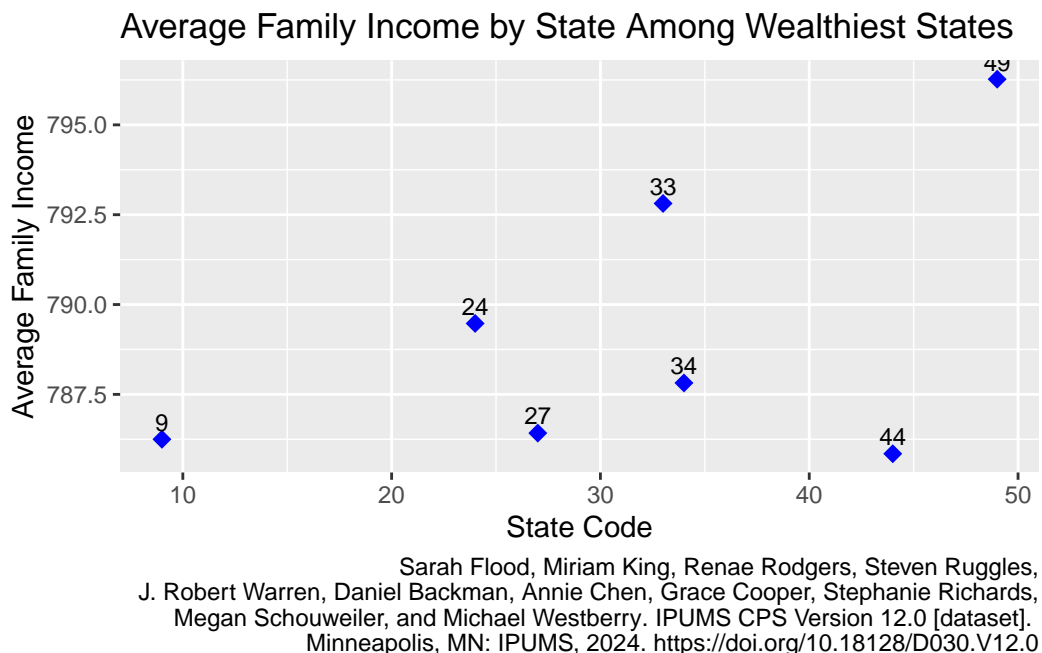
We then grew curious of how the wealthiest states fared in terms of food security scores, so we decided to identify the wealthiest states so we could analyze them further. This graph identifies the 7 wealthiest states based on average family income. Each point has the number state on top of it, so we can identify which 7 states are the wealthiest and use their FIPS codes to inform our next visualization.

Graph 3: Food Security by the Wealthiest States

```r
#Data Visualization 3:
rich_state_codes <- c(9, 24, 27, 33, 34, 44, 49)
avg_fsrawscra <- data |>
  group_by(STATEFIP) |>
  summarize(num_10 = sum(FSRAWSCRA == 10, na.rm = TRUE)) |>
  summarize(avg_num_10 = mean(num_10)) |>
  pull(avg_num_10)

data |>
filter(FSRAWSCRA == 10, STATEFIP %in% rich_state_codes) |>
  mutate(STATEFIP = as.numeric(STATEFIP)) |>
  ggplot(aes(x = factor(STATEFIP), y = FSRAWSCRA, alpha = 0.2)) +
geom_col(fill="pink") +
  geom_hline(yintercept = avg_fsrawscra, linetype = "dotted",
  color = "magenta", size = 1) +
  labs(
    title = "Number of Respondents in the 7 Wealthiest States with
    a Food Security Score of 10",
    x = "State Code",
    y = "Food Security Score",
     caption = "Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles,
J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards,
Megan Schouweiler, and Michael Westberry. IPUMS CPS Version 12.0 [dataset].
Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0"
    )
```

Number of Respondents in the 7 Wealthiest States with a Food Security Score of 10

Once we identified the 7 wealthiest states by average family income, we used this graph to see how many of each state's respondents had a food security score of 10, the highest score. By comparison, the dotted line shows the average number of respondents per state with a food security score of 10 across all states in the US.

Graph 4: Calculated Percentage of Respondents

```r
#Total respondents by state
tibble1 <- data |>
count(STATEFIP)

#Scores of 10 by State
tibble2 <- data |>
filter(FSRAWSCRA == 10) |>
count(STATEFIP)

#Percentages of the top 7 States
percents <- c(0.002, 0.001, 0.003, 0.002, 0.001, 0.001, 0.0005)
rich_state_codes <- c(9, 24, 27, 33, 34, 44, 49)

richest_states <- data.frame(percents, rich_state_codes)

ggplot(richest_states) +
geom_line(aes(x = rich_state_codes, y = percents, color = percents)) +
geom_text(aes(x = rich_state_codes, y = percents,
      label = scales::percent(percents)), vjust = -0.5, size = 3) +
  labs(
    title = "Percent of Respondents with Food Security Score = 10",
    subtitle = "Among 7 Wealthiest States",
    caption = "Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles,
J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards,
Megan Schouweiler, and Michael Westberry. IPUMS CPS Version 12.0 [dataset].
Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0"
  )+
    scale_x_continuous(breaks = rich_state_codes,
    labels = c("CT", "MD", "MN", "NH", "NJ", "RI", "UT"),
    name =  "States by their FIPS Code") +
  scale_y_continuous(name = "Percentage of Respondents")
```
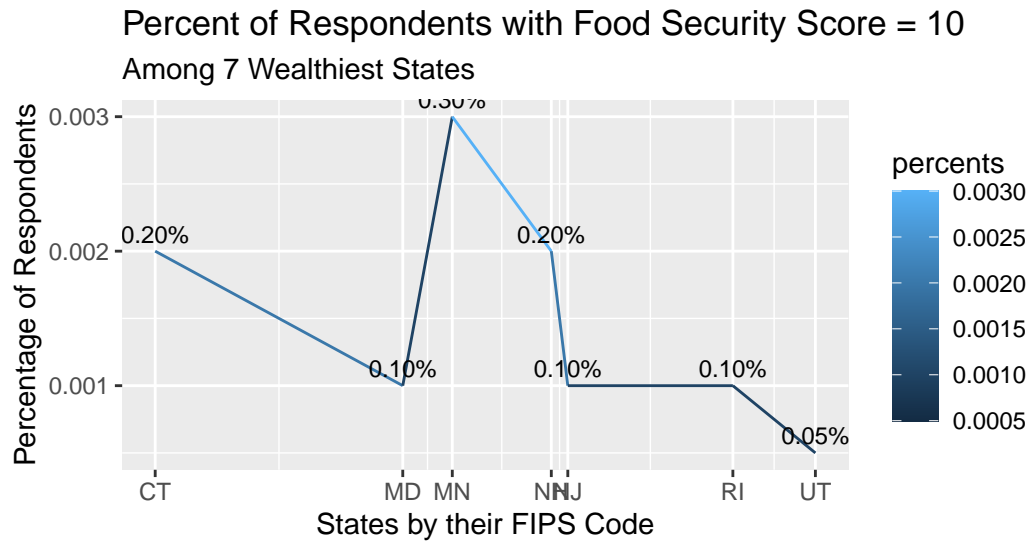
Percent of Respondents with Food Security Score = 10

Among 7 Wealthiest States

While the number of respondents per state with a food security score of 10 (displayed in the previous graph) is a helpful metric, different states have different amounts of respondents, so the comparison is harder to gauge. Thus, in the line graph above, we calculated the percentage of respondents per state that had a food security score of 10. We also changed the states' FIPS codes into state name acronyms for identification. We can see that Minnesota has the largest percentage of resondents with a food security score of 10, at 0.3%. Utah has the lowest percentage of respondents with a food security score of 10, which aligns with it being the state with the highest average household income. We did notice, however, that the 7 wealthiest states appeared to have a lower number of respondents compared to other states. This potential caveat would be worth exploring if we were to continue this analysis.