Income Bilingual Survey Design

2024-07-30

Loading in Data

```
#all crosswalks
ddi <- read_ipums_ddi("usa_00009.xml")
all_indicator_data <- read_ipums_micro(ddi)
```

```
## Use of data from IPUMS USA is subject to conditions including that users should cite
the data appropriately. Use command `ipums_conditions()` for more details.
```

```
#location data
regions <- read.csv("../location_data/County_12_Regions.csv")
rural_urban <-read.csv("../location_data/rural_urban.csv")

#language micro with bilungual
language_micro_data <- read.csv("Microdata_Bilingualism.csv")
```

Setting up Income Data

```
income_data <- language_micro_data |>
  select(AGE, INCTOT, Bilingual, PERWT, SEX, EDUCD, AGE, CLUSTER, STRATA, YEAR) |>
  filter(INCTOT != 9999999 & INCTOT > 0) |>
  filter(AGE > 14)

sum_weights_filtered <- sum(income_data$PERWT, na.rm = TRUE)

sum_weights_original <- sum(language_micro_data$PERWT, na.rm = TRUE)

#recalibrate weights
income_data <- income_data |>
  mutate(recalibrated_weight = PERWT * (sum_weights_filtered / sum_weights_original))

income_data_weighted <- income_data |>
  mutate(Weighted_Income = INCTOT * (recalibrated_weight / 100))
```

```
write.csv(file = "Income_Bilingualism_Weighted_Data.csv", income_data_weighted)
```

Using Survey Design

Previously survey design not used, so using it properly now in this methodology.

```
#setting it up
options(survey.lonely.psu = "adjust")

#creating the survey object
des <- svydesign(ids = ~1,
                 strata = ~STRATA,
                 weights = ~PERWT,
                 data = income_data)

#calculatng median by setting quantile to %50
median_income <- svyby(~INCTOT, ~Bilingual, des, svyquantile, quantiles = 0.5, ci = TRU
E)

print(median_income)
```

```
##                   Bilingual INCTOT se.INCTOT
## Bilingual         Bilingual  33994 115.56267
## Not Bilingual Not Bilingual  40000  53.57212
```

```
#splitting into proper quintiles
income_data <- income_data |>
  group_by(Bilingual)|>
  mutate(quintile = ntile(INCTOT, 5)) |>
  ungroup()

#creating a survey design with modified income data
des <- svydesign(ids = ~1,
                 strata = ~STRATA,
                 weights = ~PERWT,
                 data = income_data)

#median for each quintile in income group
median_by_quintile <- svyby(~INCTOT, ~interaction(Bilingual, quintile), des, svyquantil
e, quantiles = 0.5, ci = TRUE)

print(median_by_quintile)
```

```
##                 interaction(Bilingual, quintile) INCTOT se.INCTOT
## Bilingual.1                         Bilingual.1   6315  72.70328
## Not Bilingual.1                 Not Bilingual.1   6799  22.95898
## Bilingual.2                         Bilingual.2  18371  29.33641
## Not Bilingual.2                 Not Bilingual.2  21012  46.17305
## Bilingual.3                         Bilingual.3  32839  67.85640
## Not Bilingual.3                 Not Bilingual.3  39660  52.29545
## Bilingual.4                         Bilingual.4  54046  88.00924
## Not Bilingual.4                 Not Bilingual.4  64963  36.98946
## Bilingual.5                         Bilingual.5 103768 423.46475
## Not Bilingual.5                 Not Bilingual.5 125779 401.78210
```

```
#changing the names of columns for clarity and separating interaction
final_median_by_quintile <- median_by_quintile |>
  separate(`interaction(Bilingual, quintile)`, into = c("Bilingual", "Quintile"), sep =
"\\.") |>
  rename(median_income = INCTOT)
```

Income Medians by Education

CODES:

CODE	Educational Level
001 & 999	N/a & missing
002	no schooling completed
10-61	grade school
062	high school or GED
65-100	one or more years of college, no degree
101	bachelors
114	masters
115	professional degree beyond bachelors
116	doctoral

```
#clarity for the education codes
map_educational_level <- function(code) {
  if (code %in% c(1, 999)) {
    return("N/A or Missing")
  } else if (code == 2) {
    return("No Schooling Completed")
  } else if (code >= 10 & code < 62) {
    return("No High School Degree or GED")
  } else if (code >= 62 & code < 65) {
    return("High School or GED")
  } else if (code >= 65 & code <= 100) {
    return("Some College, No Degree")
  } else if (code == 101) {
    return("Bachelor's")
  } else if (code == 114) {
    return("Master's")
  } else if (code == 115) {
    return("Professional Degree Beyond Bachelor's")
  } else if (code == 116) {
    return("Doctoral")
  } else {
    return(NA)
  }
}

income_data_education <- income_data |>
  mutate(Educational_Level = sapply(EDUCD, map_educational_level))


income_data_education |>
  filter(is.na(Educational_Level))
```

```
## # A tibble: 0 × 12
## # ℹ 12 variables: AGE <int>, INCTOT <dbl>, Bilingual <chr>, PERWT <int>,
## #   SEX <int>, EDUCD <int>, CLUSTER <dbl>, STRATA <int>, YEAR <int>,
## #   recalibrated_weight <dbl>, quintile <int>, Educational_Level <chr>
```

```
#creating survey design for education dataset
des_edu <- svydesign(ids = ~1,
                     strata = ~STRATA,
                     weights = ~PERWT,
                     data = income_data_education)

#calculating the median income for each education level within each bilingual group
median_by_education <- svyby(~INCTOT, ~interaction(Bilingual, Educational_Level), des_ed
u, svyquantile, quantiles = 0.5, ci = TRUE)

#for clarity
median_by_education <- median_by_education |>
  separate(`interaction(Bilingual, Educational_Level)`, into = c("Bilingual", "Education
al_Level"), sep = "\\.")|>
  rename(median_income = INCTOT) |>
  rename(se_median_income = se.INCTOT) |>
  select(Bilingual, Educational_Level, median_income, se_median_income)

print(median_by_education)
```

```
##                                                    Bilingual
## Bilingual.Bachelor's                               Bilingual
## Not Bilingual.Bachelor's                       Not Bilingual
## Bilingual.Doctoral                                 Bilingual
## Not Bilingual.Doctoral                         Not Bilingual
## Bilingual.High School or GED                       Bilingual
## Not Bilingual.High School or GED               Not Bilingual
## Bilingual.Master's                                 Bilingual
## Not Bilingual.Master's                         Not Bilingual
## Bilingual.No High School Degree or GED             Bilingual
## Not Bilingual.No High School Degree or GED     Not Bilingual
## Bilingual.No Schooling Completed                   Bilingual
## Not Bilingual.No Schooling Completed           Not Bilingual
## Bilingual.Professional Degree Beyond Bachelor's    Bilingual
## Not Bilingual.Professional Degree Beyond Bachelor's Not Bilingual
## Bilingual.Some College, No Degree                  Bilingual
## Not Bilingual.Some College, No Degree          Not Bilingual
##                                                                                  Educational_L
evel
## Bilingual.Bachelor's                                                                   Bachel
or's
## Not Bilingual.Bachelor's                                                               Bachel
or's
## Bilingual.Doctoral                                                                       Doct
oral
## Not Bilingual.Doctoral                                                                   Doct
oral
## Bilingual.High School or GED                                                    High School or
GED
## Not Bilingual.High School or GED                                                High School or
GED
## Bilingual.Master's                                                                     Mast
er's
## Not Bilingual.Master's                                                                 Mast
er's
## Bilingual.No High School Degree or GED                          No High School Degree or
GED
## Not Bilingual.No High School Degree or GED                      No High School Degree or
GED
## Bilingual.No Schooling Completed                                      No Schooling Compl
eted
## Not Bilingual.No Schooling Completed                                  No Schooling Compl
eted
## Bilingual.Professional Degree Beyond Bachelor's     Professional Degree Beyond Bachel
or's
## Not Bilingual.Professional Degree Beyond Bachelor's Professional Degree Beyond Bachel
or's
## Bilingual.Some College, No Degree                                   Some College, No De
gree
## Not Bilingual.Some College, No Degree                               Some College, No De
gree
##                                                    median_income
```

```
## Bilingual.Bachelor's                                        54046
## Not Bilingual.Bachelor's                                    63000
## Bilingual.Doctoral                                          90651
## Not Bilingual.Doctoral                                      95302
## Bilingual.High School or GED                                28057
## Not Bilingual.High School or GED                            27280
## Bilingual.Master's                                          76931
## Not Bilingual.Master's                                      72131
## Bilingual.No High School Degree or GED                      21190
## Not Bilingual.No High School Degree or GED                  12500
## Bilingual.No Schooling Completed                            25000
## Not Bilingual.No Schooling Completed                        18237
## Bilingual.Professional Degree Beyond Bachelor's             91857
## Not Bilingual.Professional Degree Beyond Bachelor's        110803
## Bilingual.Some College, No Degree                           31200
## Not Bilingual.Some College, No Degree                       35670
##                                                    se_median_income
## Bilingual.Bachelor's                                       255.09814
## Not Bilingual.Bachelor's                                   289.02900
## Bilingual.Doctoral                                        2047.03437
## Not Bilingual.Doctoral                                    2158.01490
## Bilingual.High School or GED                                83.92815
## Not Bilingual.High School or GED                           159.69325
## Bilingual.Master's                                         932.60767
## Not Bilingual.Master's                                     727.26504
## Bilingual.No High School Degree or GED                     220.40358
## Not Bilingual.No High School Degree or GED                 247.68969
## Bilingual.No Schooling Completed                           612.11625
## Not Bilingual.No Schooling Completed                      1053.64437
## Bilingual.Professional Degree Beyond Bachelor's           3573.86372
## Not Bilingual.Professional Degree Beyond Bachelor's       1985.71286
## Bilingual.Some College, No Degree                          320.15209
## Not Bilingual.Some College, No Degree                      303.31672
```

CSV FILES

```
write.csv(file = "Income_SD_Median_Income.csv", median_income)
write.csv(file = "Income_SD_Quintile_Medians.csv", final_median_by_quintile)
write.csv(file = "Income_SD_ByEducation_Medians.csv", median_by_education)
```