Bilangual Populations - Percent Point Change

2024-07-23

Loading in Data

Data being used is from 5 Year Estimates from ACS of 2012 and 2022. Specifically this overview focuses on Table S1601

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

#2022
poverty_2022 <- read.csv("../ACS_DATA/2022/ACSDT5Y2022.B16009-Data.csv")
language_2022 <- read.csv("../ACS_DATA/2022/ACSST5Y2022.S1601-Data.csv")
social_2022 <- read.csv("../ACS_DATA/2022/ACSCP5Y2022.CP02-Data.csv")
characteristics_2022 <- read.csv("../ACS_DATA/2022/ACSST5Y2022.S1603-Data.csv")
limited_eng_2022 <- read.csv("../ACS_DATA/2022/ACSST5Y2022.S1602-Data.csv")
household_2022 <- read.csv("../ACS_DATA/2022/ACSDT5Y2022.B16002-Data.csv")
education_2022 <- read.csv("../ACS_DATA/2022/ACSDT5Y2022.B16010-Data.csv")

#2012
language_2012 <- read.csv("../ACS_DATA/2012/ACSST5Y2012.S1601-Data.csv")

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

CUMULATIVE TRENDS

BILANGUALISM 2012-2022

Overall methodology: Determine bilingualism by seeing proportion of each language other than Enlgish category that speaks English ‘very well’. Will then use this population in further aggregation and trend analyses.

```
bilangualism_2012 <- language_2012 |>
  select(Geographic.Area.Name,Total..Estimate..Population.5.years.and.over,
         Total..Estimate..Speak.a.language.other.than.English,
         Percent.of.specified.language.speakers..Speak.English.very.well...Estimate..Spe
ak.a.language.other.than.English.,

         #Selecting percents
         Percent.of.specified.language.speakers..Speak.English.very.well...Estimate..Spe
ak.a.language.other.than.English..Spanish.or.Spanish.Creole.,
         Percent.of.specified.language.speakers..Speak.English.very.well...Estimate..Spe
ak.a.language.other.than.English..Asian.and.Pacific.Island.languages.,
         Percent.of.specified.language.speakers..Speak.English.very.well...Estimate..Spe
ak.a.language.other.than.English..Other.Indo.European.languages.,
         Percent.of.specified.language.speakers..Speak.English..less.than.very.well...Es
timate..Speak.a.language.other.than.English..Other.languages.,

         #Selecting bilangual populations
         Total..Estimate..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Spanish.or.Spanish.Creol
e,
         Total..Estimate..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Asian.and.Pacific.Island.
languages,
         Total..Estimate..Speak.a.language.other.than.English..Other.Indo.European.langu
ages,
         Total..Estimate..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Other.languages,
         Total..Estimate..Speak.only.English
         ) |>

         #Converting to decimal for later perposes
  mutate(Total..Estimate..Speak.a.language.other.than.English = Total..Estimate..Speak.
a.language.other.than.English/ 100) |>

           #renaming for clarity cause the other names are so confusing and makes it eas
ier to just code normally later on
  rename(NonEnglish_Language_Estimate = Total..Estimate..Speak.a.language.other.than.Eng
lish,
         Spanish_Estimate = Total..Estimate..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Spanis
h.or.Spanish.Creole,
         Asian_Pacific_Estimate = Total..Estimate..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..
Asian.and.Pacific.Island.languages,
         Other_Indo_Europe_Estimate = Total..Estimate..Speak.a.language.other.than.Engli
sh..Other.Indo.European.languages,
         Other_Estimate = Total..Estimate..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Other.la
nguages,
         English_Estimate = Total..Estimate..Speak.only.English) |>

  #grabbing bilangual proportions from each language category
  rename(Percent_Bilangual = Percent.of.specified.language.speakers..Speak.English.very.
well...Estimate..Speak.a.language.other.than.English.,
       Percent_of_Spanish_Bilangual = Percent.of.specified.language.speakers..Speak.Engli
sh.very.well...Estimate..Speak.a.language.other.than.English..Spanish.or.Spanish.Creol
e.,
         Percent_of_Asian_Pacific_Bilangual = Percent.of.specified.language.speakers..Sp
```

```
eak.English.very.well...Estimate..Speak.a.language.other.than.English..Asian.and.Pacifi
c.Island.languages.,
      Percent_of_IndoEuro_Bilangual = Percent.of.specified.language.speakers..Speak.Engl
ish.very.well...Estimate..Speak.a.language.other.than.English..Other.Indo.European.langu
ages.,
      Percent_of_Other_Bilangual = Percent.of.specified.language.speakers..Speak.Englis
h..less.than.very.well...Estimate..Speak.a.language.other.than.English..Other.language
s.,
) |>

  #only catergory without numbers, so tranforming it based on percentages
  mutate(Other_Indo_Europe_Estimate = round(Total..Estimate..Population.5.years.and.over
* (Other_Indo_Europe_Estimate /100), 0)) |>

  #making blank values 0
  mutate(Percent_of_Spanish_Bilangual = ifelse(Percent_of_Spanish_Bilangual == "-", 0, a
s.numeric(Percent_of_Spanish_Bilangual)),
         Percent_of_Asian_Pacific_Bilangual = ifelse(Percent_of_Asian_Pacific_Bilangual
== "-", 0, as.numeric(Percent_of_Asian_Pacific_Bilangual)),
         Percent_of_IndoEuro_Bilangual = ifelse(Percent_of_IndoEuro_Bilangual == "-", 0,
as.numeric(Percent_of_IndoEuro_Bilangual)),
         Percent_of_Other_Bilangual = ifelse(Percent_of_Other_Bilangual == "-", 0, as.nu
meric(Percent_of_Other_Bilangual)),
         Percent_Bilangual = ifelse(Percent_Bilangual == "-", 0, as.numeric(Percent_Bila
ngual) / 100)
         ) |>

  #creating total bilangual based on people that speak the language and multiplying by p
roportion that speak the language and speak english very well
  mutate(Spanish_Bilangual = round(Spanish_Estimate * (Percent_of_Spanish_Bilangual / 10
0), 0),
         Asian_Pacific_Bilangual = round(Asian_Pacific_Estimate * (Percent_of_Asian_Paci
fic_Bilangual / 100), 0),
       IndoEuro_Bilangual = round(Other_Indo_Europe_Estimate * (Percent_of_IndoEuro_Bila
ngual / 100), 0),
       Other_Bilangual = round(Other_Estimate * (Percent_of_Other_Bilangual / 100), 0)
) |>

  mutate(Percent_Spanish_Bilangual = round((Spanish_Bilangual / Total..Estimate..Populat
ion.5.years.and.over), 3),
         Percent_Asian_Pacific_Bilangual  = round((Asian_Pacific_Bilangual/ Total..Estim
ate..Population.5.years.and.over), 3),
          Percent_IndoEuro_Bilangual = round((IndoEuro_Bilangual / Total..Estimate..Popu
lation.5.years.and.over), 3),
         Percent_Other_Bilangual = round((Other_Bilangual / Total..Estimate..Population.
5.years.and.over), 3)

)
```

```
## Warning: There were 5 warnings in `mutate()`.
## The first warning was:
## ℹ In argument: `Percent_of_Spanish_Bilangual =
##   ifelse(Percent_of_Spanish_Bilangual == "–", 0,
##   as.numeric(Percent_of_Spanish_Bilangual))`.
## Caused by warning in `ifelse()`:
## ! NAs introduced by coercion
## ℹ Run `dplyr::last_dplyr_warnings()` to see the 4 remaining warnings.
```

```
write.csv(file = "FULL_Bilingual_2012.csv", bilangualism_2012)
```

```
bilangualism_2022 <- language_2022 |>
  select(Geographic.Area.Name, Estimate..Total..Population.5.years.and.over,
         Estimate..Total..Population.5.years.and.over..Speak.a.language.other.than.Engli
sh,

         #grabbing people who speak the language proportions from each language category
         Estimate..Total..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLI
SH..Spanish,
         Estimate..Total..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLI
SH..Asian.and.Pacific.Island.languages,
         Estimate..Total..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLI
SH..Other.Indo.European.languages,
         Estimate..Total..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLI
SH..Other.languages,
         Estimate..Total..Population.5.years.and.over..Speak.only.English,

         #grabbing bilangual proportions from each language category

         Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.
language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Spa
nish.,
         Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.
language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Asi
an.and.Pacific.Island.languages.,
         Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.
language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Oth
er.Indo.European.languages.,
         Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.
language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Oth
er.languages.,
         Estimate..Speak.English..less.than.very.well...Percent.of.specified.language.sp
eakers..Population.5.years.and.over..Speak.a.language.other.than.English.,
         Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.
language.speakers..Population.5.years.and.over..Speak.a.language.other.than.English.) |>

           #renaming for clarity

  rename(NonEnglish_Language_Estimate = Estimate..Total..Population.5.years.and.over..Sp
eak.a.language.other.than.English,
         Spanish_Estimate = Estimate..Total..Population.5.years.and.over..SPEAK.A.LANGUA
GE.OTHER.THAN.ENGLISH..Spanish,
         Asian_Pacific_Estimate = Estimate..Total..Population.5.years.and.over..SPEAK.A.
LANGUAGE.OTHER.THAN.ENGLISH..Asian.and.Pacific.Island.languages,
         Other_Indo_Europe_Estimate = Estimate..Total..Population.5.years.and.over..SPEA
K.A.LANGUAGE.OTHER.THAN.ENGLISH..Other.Indo.European.languages,
         Other_Estimate = Estimate..Total..Population.5.years.and.over..SPEAK.A.LANGUAG
E.OTHER.THAN.ENGLISH..Other.languages,



         Spanish_Bilangual = Estimate..Speak.English.only.or.speak.English.very.well...P
ercent.of.specified.language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTH
ER.THAN.ENGLISH..Spanish.,
```

```
         Asian_Pacific_Bilangual = Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Asian.and.Pacific.Island.languages.,
      IndoEuro_Bilangual = Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Other.Indo.European.languages.,
      Other_Bilangual = Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.language.speakers..Population.5.years.and.over..SPEAK.A.LANGUAGE.OTHER.THAN.ENGLISH..Other.languages.,
      Other_Langual_Not_Bilingual = Estimate..Speak.English..less.than.very.well...Percent.of.specified.language.speakers..Population.5.years.and.over..Speak.a.language.other.than.English.,
      English_Monolangual = Estimate..Total..Population.5.years.and.over..Speak.only.English,
      Overall_Bilingual = Estimate..Speak.English.only.or.speak.English.very.well...Percent.of.specified.language.speakers..Population.5.years.and.over..Speak.a.language.other.than.English.) |>


  #creating total bilangual based on people that speak the language and multiplying by proportion that speak the language and speak english very well
    mutate(
          Percent_of_Spanish_Bilangual = ifelse(Spanish_Estimate == 0 | Spanish_Bilangual == 0, 0, round(Spanish_Bilangual / Spanish_Estimate, 3)),
          Percent_of_Asian_Pacific_Bilangual = ifelse(Asian_Pacific_Estimate == 0 | Asian_Pacific_Bilangual == 0, 0, round(Asian_Pacific_Bilangual / Asian_Pacific_Estimate, 3)),
          Percent_of_IndoEuro_Bilangual = ifelse(IndoEuro_Bilangual == 0 | Other_Indo_Europe_Estimate == 0, 0, round(IndoEuro_Bilangual / Other_Indo_Europe_Estimate, 3)),
          Percent_of_Other_Bilangual = ifelse(Other_Bilangual == 0 | Other_Estimate == 0, 0, round(Other_Bilangual / Other_Estimate, 3)),
          Percent_Overall_Bilingual = round((Overall_Bilingual / Estimate..Total..Population.5.years.and.over), 3)
  ) |>

   mutate(
          Percent_Spanish_Bilangual = ifelse(Spanish_Estimate == 0 | Spanish_Bilangual == 0, 0, round(Spanish_Bilangual / Estimate..Total..Population.5.years.and.over, 3)),
          Percent_Asian_Pacific_Bilangual = ifelse(Asian_Pacific_Estimate == 0 | Asian_Pacific_Bilangual == 0, 0, round(Asian_Pacific_Bilangual / Estimate..Total..Population.5.years.and.over, 3)),
          Percent_IndoEuro_Bilangual = ifelse(IndoEuro_Bilangual == 0 | Other_Indo_Europe_Estimate == 0, 0, round(IndoEuro_Bilangual / Estimate..Total..Population.5.years.and.over, 3)),
          Percent_Other_Bilangual = ifelse(Other_Bilangual == 0 | Other_Estimate == 0, 0, round(Other_Bilangual / Estimate..Total..Population.5.years.and.over, 3))
  )
```

```
write.csv(file = "FULL_Bilingual_2022.csv", bilangualism_2022)
```

```
final_bilangual_2012 <- bilangualism_2012 |>
  select(Geographic.Area.Name, Total..Estimate..Population.5.years.and.over, Spanish_Bil
angual, Asian_Pacific_Bilangual, IndoEuro_Bilangual, Other_Bilangual, Percent_Spanish_Bi
langual, Percent_Asian_Pacific_Bilangual, Percent_IndoEuro_Bilangual, Percent_Other_Bila
ngual) |>
  rename(County = Geographic.Area.Name,
         Population_5_Years_Over = Total..Estimate..Population.5.years.and.over) |>
  mutate(County = sub(" County, Texas", "", County))
```

```
final_bilangual_2012 <- left_join(final_bilangual_2012, regions, by = "County")
final_bilangual_2012 <- left_join(final_bilangual_2012, rural_urban, by = "County")
```

```
final_bilangual_2022 <- bilangualism_2022 |>
  select(Geographic.Area.Name, Estimate..Total..Population.5.years.and.over, Spanish_Bil
angual, Asian_Pacific_Bilangual, IndoEuro_Bilangual, Other_Bilangual, Percent_Spanish_Bi
langual, Percent_Asian_Pacific_Bilangual, Percent_IndoEuro_Bilangual, Percent_Other_Bila
ngual) |>
  rename(County = Geographic.Area.Name,
         Population_5_Years_Over = Estimate..Total..Population.5.years.and.over) |>
  mutate(County = sub(" County, Texas", "", County))
```

```
final_bilangual_2022 <- left_join(final_bilangual_2022, regions, by = "County")
final_bilangual_2022 <- left_join(final_bilangual_2022, rural_urban, by = "County")
```

```
write.csv(file = "FINAL_Bilingual_2012.csv", final_bilangual_2012)
write.csv(file = "FINAL_Bilingual_2022.csv", final_bilangual_2022)
```

RURAL/URBAN

Overall methodology: Creating weights for each county based on its proportion of population of people 5 and older compared to the whole state's population of 5 and over. After this weight was multiplied by original number, aggregated by rural and urban status to compose final aggregate.

```
urban_rural_totals <- function(bilangual_data) {
  rural_urban_tot <- bilangual_data |>
  group_by(Rural_Urban_Stat)|>
  summarise(Region_Total = sum(Population_5_Years_Over, na.rm = TRUE))

aggregated_precursor <- bilangual_data|>
  left_join(rural_urban_tot, by = "Rural_Urban_Stat") |>
  mutate(Weight = Population_5_Years_Over / Region_Total)

aggregated_data <- aggregated_precursor |>
  group_by(Rural_Urban_Stat) |>
  summarise(
    Total_Spanish_Bilangual = sum(Spanish_Bilangual * Weight, na.rm = TRUE),
    Total_Asian_Pacific_Bilangual = sum(Asian_Pacific_Bilangual * Weight, na.rm = TRUE),
    Total_IndoEuro_Bilangual = sum(IndoEuro_Bilangual * Weight, na.rm = TRUE),
    Total_Other_Bilangual = sum(Other_Bilangual * Weight, na.rm = TRUE),
    Total_Population_5_Years_Over = sum(Population_5_Years_Over * Weight, na.rm = TRUE)
  ) |>
  mutate(
    Percent_Spanish_Bilangual = round(Total_Spanish_Bilangual / Total_Population_5_Years
_Over, 3),
    Percent_Asian_Pacific_Bilangual = round(Total_Asian_Pacific_Bilangual / Total_Popula
tion_5_Years_Over, 3),
    Percent_IndoEuro_Bilangual = round(Total_IndoEuro_Bilangual / Total_Population_5_Yea
rs_Over, 3),
    Percent_Other_Bilangual = round(Total_Other_Bilangual / Total_Population_5_Years_Ove
r, 3)
  )

return(aggregated_data)
}
```

```
rural_urban_2012 <- urban_rural_totals(final_bilangual_2012)
rural_urban_2022 <- urban_rural_totals(final_bilangual_2022)
```

REGIONS

Overall methodology: Creating weights for each county based on its proportion of population of people 5 and older compared to the whole state's population of 5 and over. After this weight was multiplied by original number, aggregated by region category to compose final aggregate.

```
region_total <- function(bilangual_data) {
  region_totals <- bilangual_data |>
  group_by(Region)|>
  summarise(Region_Total = sum(Population_5_Years_Over, na.rm = TRUE))

aggregated_precursor <- bilangual_data|>
  left_join(region_totals, by = "Region") |>
  mutate(Weight = Population_5_Years_Over / Region_Total)

aggregated_data <- aggregated_precursor |>
  group_by(Region) |>
  summarise(
    Total_Spanish_Bilangual = sum(Spanish_Bilangual * Weight, na.rm = TRUE),
    Total_Asian_Pacific_Bilangual = sum(Asian_Pacific_Bilangual * Weight, na.rm = TRUE),
    Total_IndoEuro_Bilangual = sum(IndoEuro_Bilangual * Weight, na.rm = TRUE),
    Total_Other_Bilangual = sum(Other_Bilangual * Weight, na.rm = TRUE),
    Total_Population_5_Years_Over = sum(Population_5_Years_Over * Weight, na.rm = TRUE)
  ) |>
  mutate(
    Percent_Spanish_Bilangual = round(Total_Spanish_Bilangual / Total_Population_5_Years
_Over, 3),
    Percent_Asian_Pacific_Bilangual = round(Total_Asian_Pacific_Bilangual / Total_Popula
tion_5_Years_Over, 3),
    Percent_IndoEuro_Bilangual = round(Total_IndoEuro_Bilangual / Total_Population_5_Yea
rs_Over, 3),
    Percent_Other_Bilangual = round(Total_Other_Bilangual / Total_Population_5_Years_Ove
r, 3)
  )

return(aggregated_data)
}
```

```
regions_2012 <- region_total(final_bilangual_2012)
regions_2022 <- region_total(final_bilangual_2022)
```

```
write.csv(file = "Bilingual_Regions_2012.csv",regions_2012 )
write.csv(file = "Bilingual_Regions_2022.csv", regions_2022)
```

PERCENT POINT CHANGES

```
to_rur_urb_merge_2012 <- rural_urban_2012 |>
  select(Percent_Spanish_Bilangual, Percent_Asian_Pacific_Bilangual, Percent_IndoEuro_Bi
langual, Percent_Other_Bilangual, Rural_Urban_Stat) |>
  filter(Rural_Urban_Stat != "State") |>
  rename(Percent_Spanish_Bilangual_2012 = Percent_Spanish_Bilangual,
         Percent_Asian_Pacific_Bilangual_2012 = Percent_Asian_Pacific_Bilangual,
         Percent_IndoEuro_Bilangual_2012 = Percent_IndoEuro_Bilangual,
         Percent_Other_Bilangual_2012 = Percent_Other_Bilangual
         )

to_rur_urb_merge_2022 <- rural_urban_2022 |>
  select(Percent_Spanish_Bilangual, Percent_Asian_Pacific_Bilangual, Percent_IndoEuro_Bi
langual, Percent_Other_Bilangual, Rural_Urban_Stat) |>
  filter(Rural_Urban_Stat != "State")

differentials_rural_urban <- merge(to_rur_urb_merge_2012, to_rur_urb_merge_2022, by = "R
ural_Urban_Stat") |>
  mutate(Spanish = (Percent_Spanish_Bilangual - Percent_Spanish_Bilangual_2012) * 100,
         Asian_Pacific_Island_Languages = (Percent_Asian_Pacific_Bilangual - Percent_Asi
an_Pacific_Bilangual_2012) * 100,
         Other_Indo_European_Language = (Percent_IndoEuro_Bilangual - Percent_IndoEuro_B
ilangual_2012) * 100,
         Other_Language = (Percent_Other_Bilangual - Percent_Other_Bilangual_2012) * 100
         ) |>
  select(Rural_Urban_Stat, Spanish, Asian_Pacific_Island_Languages, Other_Indo_European_
Language, Other_Language )
```

```
to_merge_regions_2012 <- regions_2012 |>
  select(Percent_Spanish_Bilangual, Percent_Asian_Pacific_Bilangual, Percent_IndoEuro_Bi
langual, Percent_Other_Bilangual, Region) |>
  filter(Region != "State") |>
  rename(Percent_Spanish_Bilangual_2012 = Percent_Spanish_Bilangual,
         Percent_Asian_Pacific_Bilangual_2012 = Percent_Asian_Pacific_Bilangual,
         Percent_IndoEuro_Bilangual_2012 = Percent_IndoEuro_Bilangual,
         Percent_Other_Bilangual_2012 = Percent_Other_Bilangual
         )

to_merge_regions_2022 <- regions_2022 |>
  select(Percent_Spanish_Bilangual, Percent_Asian_Pacific_Bilangual, Percent_IndoEuro_Bi
langual, Percent_Other_Bilangual, Region) |>
  filter(Region != "State")

differentials_regions <- merge(to_merge_regions_2012, to_merge_regions_2022, by = "Regio
n") |>
  mutate(Spanish = (Percent_Spanish_Bilangual - Percent_Spanish_Bilangual_2012) * 100,
         Asian_Pacific_Island_Languages = (Percent_Asian_Pacific_Bilangual - Percent_Asi
an_Pacific_Bilangual_2012) * 100,
         Other_Indo_European_Language = (Percent_IndoEuro_Bilangual - Percent_IndoEuro_B
ilangual_2012) * 100,
         Other_Language = (Percent_Other_Bilangual - Percent_Other_Bilangual_2012) * 100
         ) |>
  select(Region, Spanish, Asian_Pacific_Island_Languages, Other_Indo_European_Language,
Other_Language )
```

```
write.csv(file = "Regions_Bilingual_Differential.csv", differentials_regions)
write.csv(file = "Rural_Urban_Bilingual_Differential.csv", differentials_rural_urban)
```