Analyzing the Income Disparities between Foreign-Born and Native-Born Workers in the U.S.

and Factors Affecting the Income Levels of Foreign-Born Workers

Jing Chen, Dvija Muktesh Shah, Jaini Chetan Gala

Introduction & Motivation

According to the U.S. Census Bureau ACS 5-year estimates, the median personal income of native-born workers is $38,000, and the estimated median income of foreign-born workers is $32,000. This report studies whether native-born workers are still paid more than foreign-born workers after controlling for demographic factors, educational level, work hours, and occupation.

Goals:

- Provide a better understanding of the difference in income levels between native-born and foreign-born workers in the U.S.
- Found the factors that are strongly correlated with the immigrants' income levels and provide recommendations on how to improve income levels for low-income immigrant families.
- Provide a helpful tool for people to understand better the spatial distribution of immigrants and their basic information, such as median income level and their region of birth.

Data

In this project, we use the 2005-2019 ACS 1-Year Estimates Public Use Microdata Sample (PUMS) Data which includes variables for nearly every question on the American Community survey. The ACS survey covers approximately one percent of the United States population.

Key Variable	Type	Census Definition
AGEP	Numeric	{1 to 99 years (Top-coded)}
SEX	Character	{Male & Female}
RA1CP	Character	Detailed race code
CIT	Character	Citizenship status {1 .Born in the U.S. 2 .Born in Puerto Rico, Guam, the U.S. Virgin Islands, or the .Northern Marianas 3 .Born abroad of American parent(s) 4 .U.S. citizen by naturalization 5 .Not a citizen of the U.S.}
ENG	Character	Ability to speak English b .N/A (speaks only English) 1 .Very well 2 .Well 3 .Not well 4 .Not at all
SCHL	Character	Educational attainment (See next slide for the categories)
ESR	Character	Employment status recode b .N/A (less than 16 years old) {1 .Civilian employed, at work 2 .Civilian employed, with a job but not at work 3 .Unemployed 4 .Armed forces, at work 5 .Armed forces, with a job but not at work 6 .Not in labor force}
PINCP	Character	Total person's income
POBP	Character	Place of birth
ST	Character	State
WKHP	Numeric	Usual hours worked per week past 12 months
OCCP	Character	Occupation recode for 2018 and later based on 2018 OCC codes
PUMA	Character	Public use microdata area code (PUMA) based on 2010 Census definition (areas with population of 100,000 or more)

Data Retrieving and Processing

- Load 2019 U.S Census PUMS data of all states in U.S. using get_pums() function under the tidycensus package
- Filter to only contain people who are employed, 25 years old or above, do not work from home, do not work in the military, and with wage greater than 0 (Above 1 million rows)
- Group 530 four-digit OCCP codes into nine groups and 26 Educational Attainment Code into 4 levels
- Partition data for EDA and CDA

2. Some College/Associate Degree
- Some college less than 1 year
- 1 or more year of college credits but no degree
- Associate's Degree

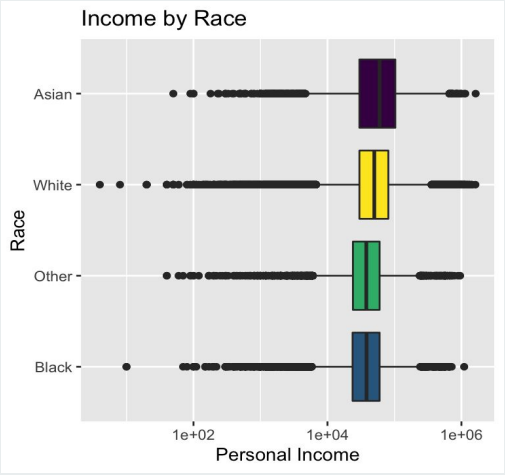
4. Less Than High School
- Null Values
- No Schooling Completed
- Nursery
- Preschool
- Kindergarten
- Grade 1-11
- No Diploma

1. Bachelor Degree or Higher
- Bachelor's Degree
- Master's Degree
- Professional Degree beyond Bachelor's Degree

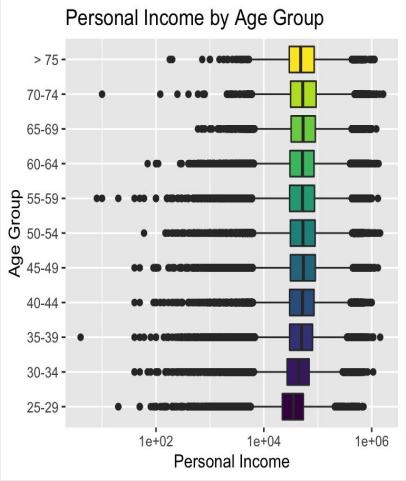
3. High School Graduate/ GED
- Regular High School
- GED or Alternative Credential

Exploratory Data Analysis

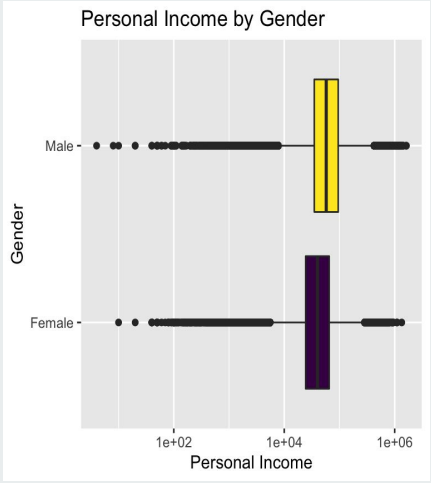
Relationship between Independent and Dependent Variables (PART 1)



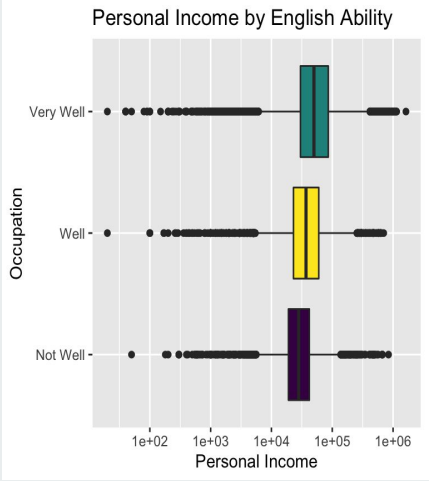
- Asian workers tend to have higher income level compared to workers of different races.



- Personal income level tend to increase as age increases.
- We see a slight decreases in personal income level after age of 60 years.



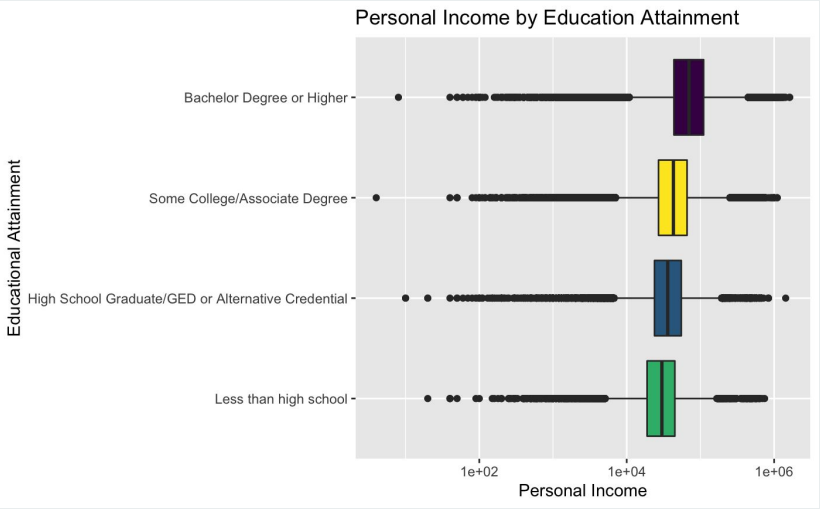
- Male workers tend to earn slightly higher compared to female workers.

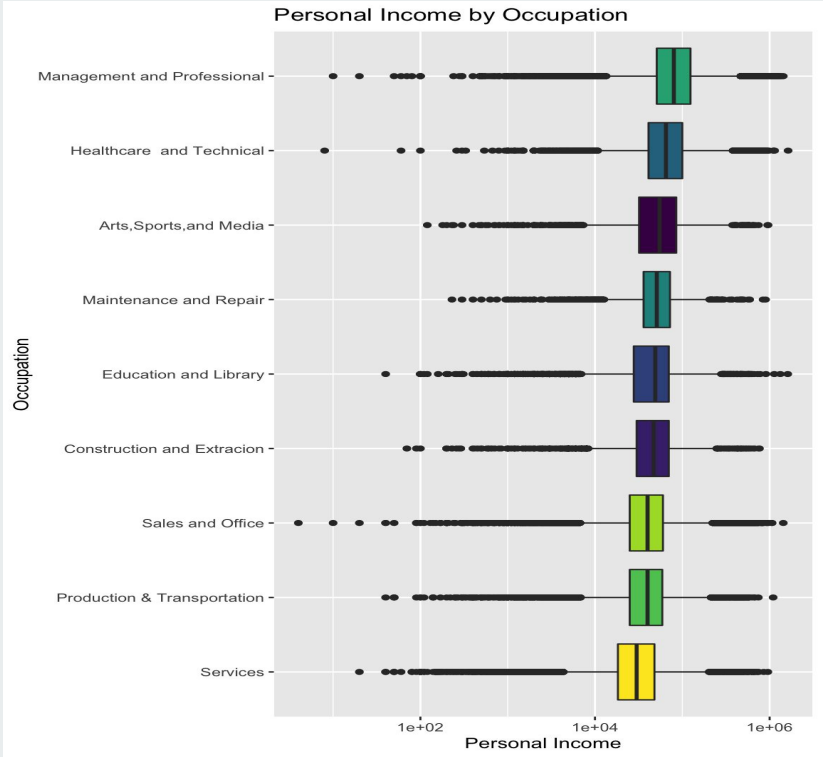


- Workers having better English speaking ability tend to earn more than other workers

Exploratory Data Analysis

Relationship between Independent and Dependent Variables (PART 2)

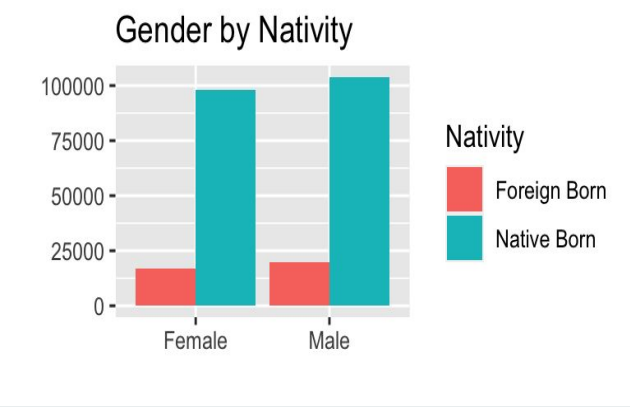


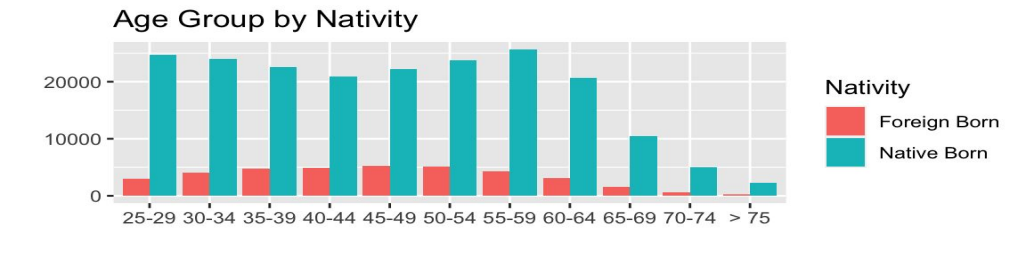


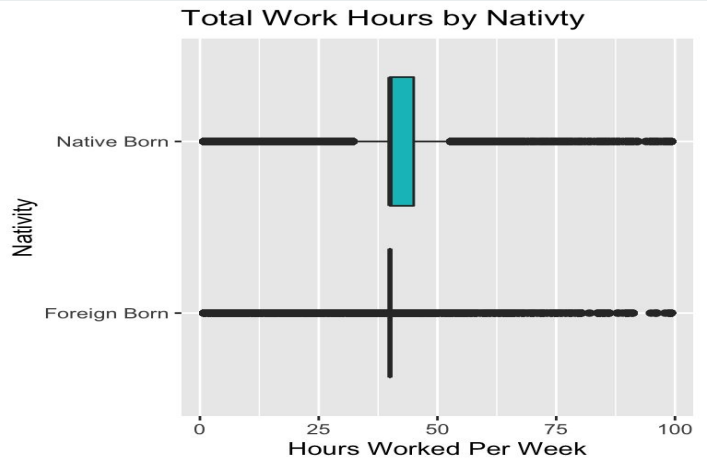
- Workers with a Bachelor Degree or Higher tend to have a higher personal income level compared to workers with other educational attainment.
- Workers working in Management and Professional or Healthcare and Technical occupation earn more than the other occupation workers.

Exploratory Data Analysis

Relationship between nativity and other Independent Variables (PART 1)



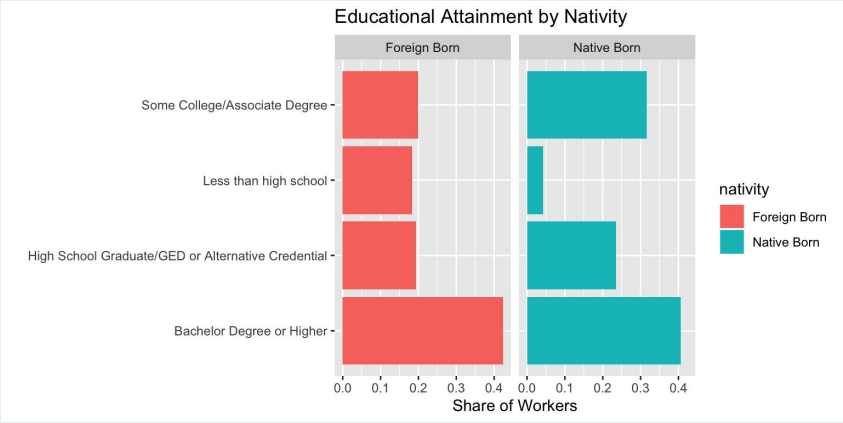


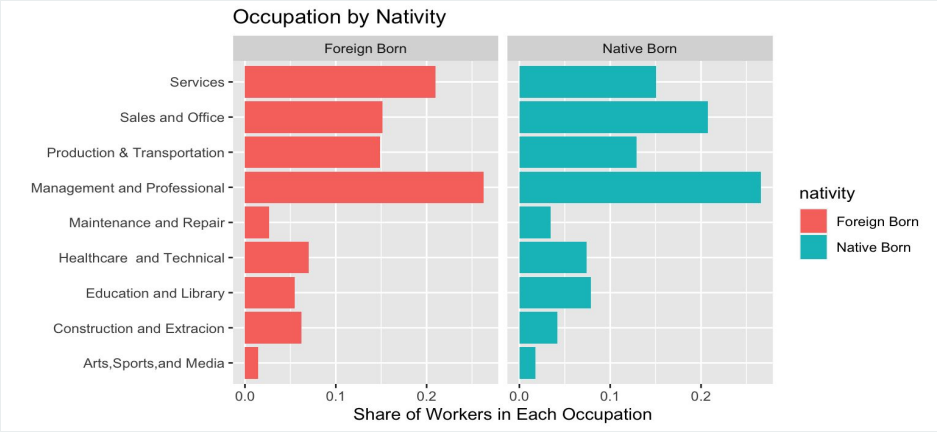


- America has higher count of native born female and native born male workers compared to foreign born female and foreign born male workers.
- The difference between the count of male and female workers is not much.
- There are large number of native born workers between the age 25-29 and 55-59 years and there are large number of foreign born workers between the age 45-54 years.
- Native born work for more hours per week compared to foreign born workers.

Exploratory Data Analysis

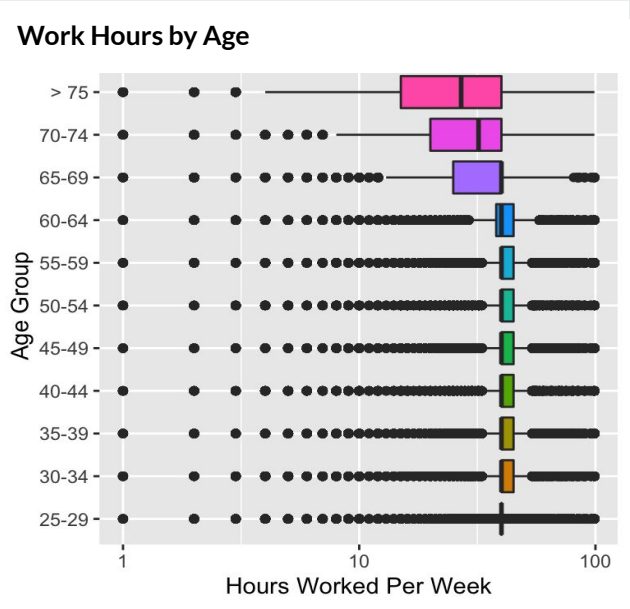
Relationship between nativity and other Independent Variables (PART 2)

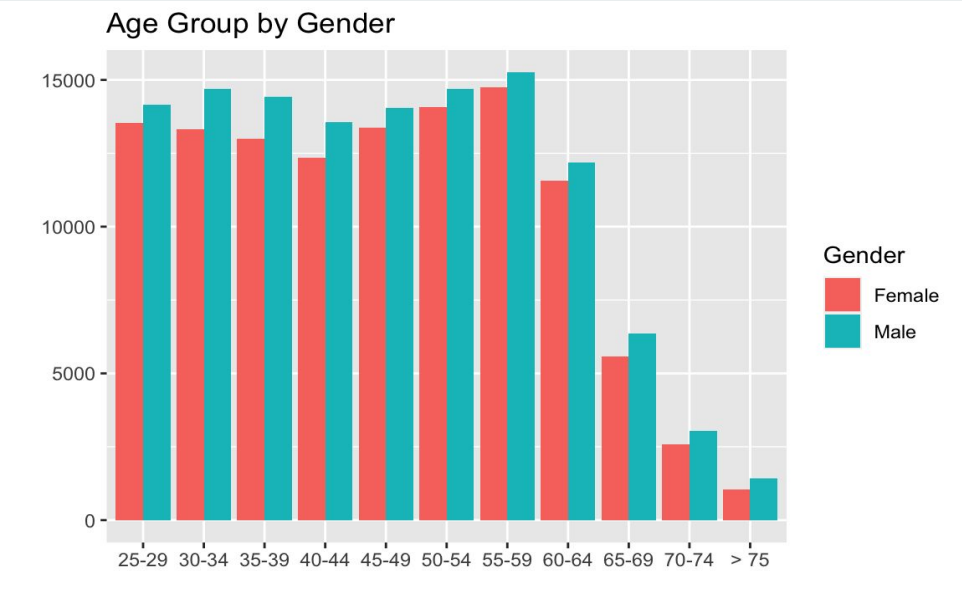




- Large number of foreign people tend to have bachelor or some college or associate degree, whereas large number of foreign people tend to have bachelors or higher degree.
- People working in Management and Professional occupation or Sales and Office tend to earn higher compared to workers having other occupations.
- There are large number of Native born workers having Management and Professional or Sales and Office as their occupation, whereas large number of foreign born workers work in Management and Professional or Service industries.

Exploratory Data Analysis





- Except for the age group 70-74 and above 75 years, all other age group has same median hours worked per week.
- Number of male workers are more compared to female workers for all age groups.
- The difference in the count of workers between male and female is consistent throughout all the age groups.

Methodology and Model

PUMS data has replicate person weights associated with each observation. We use **as_survey** and ***svyglm*** in ***survey*** package is used for building linear regression model on survey-based data with sample weight.

PWGTP1	PWGTP2	PWGTP3	PWGTP4	PWGTP5	PWGTP6	PWGTP7	PWGTP8	PWGTP9	PWGTP10	I
14	15	15	13	26	15	13	1	1	30	
150	75	78	78	7	78	79	147	7	147	
1	1	26	24	23	26	2	21	3	29	
0	13	13	11	22	11	1	12	1	2	
57	112	60	54	55	4	4	53	5	107	
39	1	22	38	37	21	18	20	23	20	
19	19	36	18	2	19	0	1	38	18	
15	1	2	29	31	0	14	30	1	1	
34	28	180	376	176	27	187	211	32	214	
115	215	118	99	113	112	198	111	123	10	
7	5	10	10	0	0	9	5	0	10	
9	0	1	4	0	6	4	9	4	11	

Model 1: Log Income = $\beta_0 + \beta_1$Nativity

Null Hypothesis (H_0): *There is no difference in personal income level between native-born and foreign-born workers*

Alternative Hypothesis (H_a): *There is a difference in personal income level between native-born and foreign-born workers*

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        10.602093   0.002788 3803.22   <2e-16 ***
nativityNative Born  0.123783   0.003086   40.11   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

- If we don’t control other factors, our team rejects the null hypothesis at the 0.001 significance level.
- There is a difference in personal income level between Native-born and Foreign-born workers
- The log Personal Income of Native-born workers is 0.12 units greater than the log Personal Income of Foreign-born workers.

Model & Results

Model 2: Log Income = β_0 + β_1Nativity + β_2 AGEP + β_3Gender + β_4Race + β_5WKHP + β_6English + α Educational Attainment Factors + γ Occupation Factors

Null Hypothesis (H_0): *There is no difference in personal income level between native-born and foreign-born workers*
Alternative Hypothesis (H_a): *There is a difference in personal income level between native-born and foreign-born workers*

```
Coefficients:
                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                              8.507e+00  1.141e-02 745.274  < 2e-16 ***
nativityNative Born                     -1.157e-02  3.122e-03  -3.705 0.000468 ***
AGEP                                     1.444e-02  7.289e-05 198.137  < 2e-16 ***
genderMale                               2.531e-01  2.347e-03 107.857  < 2e-16 ***
WKHP                                     2.793e-02  1.203e-04 232.216  < 2e-16 ***
education_tBachelor Degree or Higher     4.687e-01  2.790e-03 167.996  < 2e-16 ***
education_Less than high school         -1.218e-01  3.949e-03 -30.854  < 2e-16 ***
education_Some College/Associate Degree  1.115e-01  2.538e-03  43.926  < 2e-16 ***
english_abilityVery Well                 1.358e-01  5.262e-03  25.819  < 2e-16 ***
english_abilityNot Well                 -9.076e-02  7.219e-03 -12.571  < 2e-16 ***
raceAsian                                4.472e-02  4.381e-03  10.208 1.18e-14 ***
raceBlack                               -1.216e-01  3.172e-03 -38.336  < 2e-16 ***
raceOther                               -3.738e-02  3.576e-03 -10.452 4.79e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Reference level for Categorical Variables:
Educational Attainment (High School Degree or GED); Race (White);and English
Ability(Well).
```

- After controlling for additional variables, we still reject the null hypothesis at the 0.001 significance level.
- There is a still a difference in personal income level between native-born and foreign-born.
- However, after controlling for other factors, such as age, gender, race, the number of hours worked per week, English speaking ability, educational attainment, and occupation, the result shows that the income level for native-born is lower than foreign-born workers.

Model & Results

Model 3: Log Income = β_0 + β_1Nativity +β_2 AGEP + β_3Gender + β_4Race +β_5WKHP + β_6English + **β_7ST**
+ α Educational Attainment Factors + γ Occupation Factors

Null Hypothesis (H_0): *There is no difference in personal income level between native-born and foreign-born workers*
Alternative Hypothesis (H_a): *There is a difference in personal income level between native-born and foreign-born workers*

```
Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                       8.342e+00  1.440e-02 579.261  < 2e-16 ***
nativityNative Born               1.709e-02  3.183e-03   5.369 0.000451 ***
AGEP                              1.445e-02  7.408e-05 195.091  < 2e-16 ***
genderMale                        2.500e-01  2.330e-03 107.307 2.69e-15 ***
WKHP                              2.819e-02  1.189e-04 237.119  < 2e-16 ***
education_BachelorDegree/Higher 4.535e-01  2.678e-03 169.391  < 2e-16 ***
education_Less than high school-1.191e-01  3.909e-03 -30.466 2.16e-10 ***
education_Some College Degree   1.082e-01  2.503e-03  43.244 9.44e-12 ***
english_abilityVery Well        1.422e-01  5.245e-03  27.119 6.11e-10 ***
english_abilityNot Well        -9.022e-02  7.163e-03 -12.596 5.09e-07 ***
raceAsian                       3.634e-03  4.457e-03   0.815 0.435902
raceBlack                      -1.169e-01  3.395e-03 -34.415 7.29e-11 ***
raceOther                      -7.479e-02  3.817e-03 -19.593 1.09e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Reference level for Categorical Variables:
Educational Attainment (High School Degree or GED); Race (White);and
English Ability(Well).
```

- We still reject the null hypothesis at significance level 0.001.
- Adding one more predictor that is State to model 2 shows a difference in the personal income level between native-born workers and foreign-born workers.
- In model 2, the personal income of native-born workers is 0.0115 units lower than the personal income of foreign-born workers.
- And after adding the State predictor, it changes to 0.0170 units higher than the personal income of foreign-born workers.

Model & Takeaways

Model 4: Log Income = β_0 + β_1English +β_2 AGEP + β_3Gender + β_4Race +β_5WKHP + β_6ST + α Educational Attainment Factors + γ Occupation Factors

```
Coefficients:                Estimate Std. Error  tvalue Pr(>|t|)
(Intercept)                 8.5367580  0.0431203 197.976  < 2e-16 ***
AGEP                        0.0118354  0.0002141  55.275 9.11e-14 ***
genderMale                  0.2507487  0.0052639  47.636 4.01e-13 ***
raceAsian                   0.0165391  0.0055412   2.985 0.013697 *
raceBlack                  -0.1134013  0.0083379 -13.601 8.93e-08 ***
raceOther                  -0.0470693  0.0064063  -7.347 2.46e-05 ***
WKHP                        0.0265202  0.0003263  81.281 1.94e-15 ***
english_abilityVery Well    0.1610192  0.0059581  27.025 1.11e-10 ***
english_abilityNot Well    -0.1013082  0.0074150 -13.663 8.55e-08 ***
When using 'High School Degree/GED' as reference level
edu_Bachelor Degree or Higher 0.3221086  0.0080104  40.211 2.16e-12 ***
edu-Less than high school   -0.0579717  0.0074223  -7.811 1.45e-05 ***
edu_Some College Degree      0.0547292  0.0070706   7.740 1.57e-05 ***
When using 'Management and Professional Occupation' as reference level
occp_Maintenance and Repair  -0.4797778  0.0143415 -33.454 1.35e-11 ***
occp_Construction & Extracion-0.4855983  0.0116120 -41.819 1.47e-12 ***
occp_Production & Transport  -0.6250248  0.0089122 -70.132 8.47e-15 ***
occp_Services                -0.7322400  0.0087137 -84.034 1.39e-15 ***
occp_Education & Library     -0.5739185  0.0120991 -47.435 4.18e-13 ***
occp_Arts,Sports,and Media   -0.3845876  0.0218576 -17.595 7.48e-09 ***
occp_Healthcare&Technical     0.0311675  0.0113838   2.738 0.020909 *
occp_Sales and Office        -0.5412648  0.0092614 -58.443 5.22e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

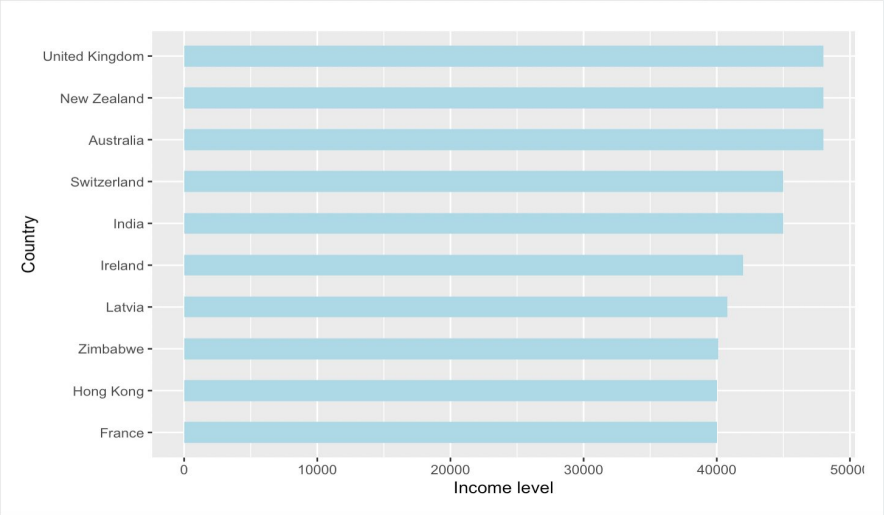
Key Takeaways

- Higher English speaking skill helps foreign-born workers to improve their income level
- Foreign-born workers with a High school degree or GED earn more than those without a High school degree.
- Foreign-born workers with an associate degree have a higher income than those with only a high school degree.
- Foreign-born workers with at least Bachelor's degree have a higher income than those with an associate degree.

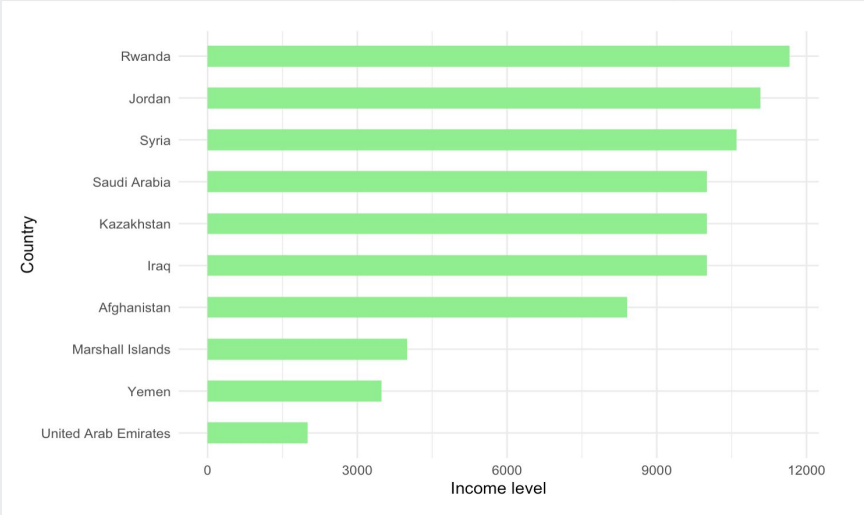
Immigrant Communities

Top 10 Immigrant communities with highest/lowest median personal income

Top 10 Highest

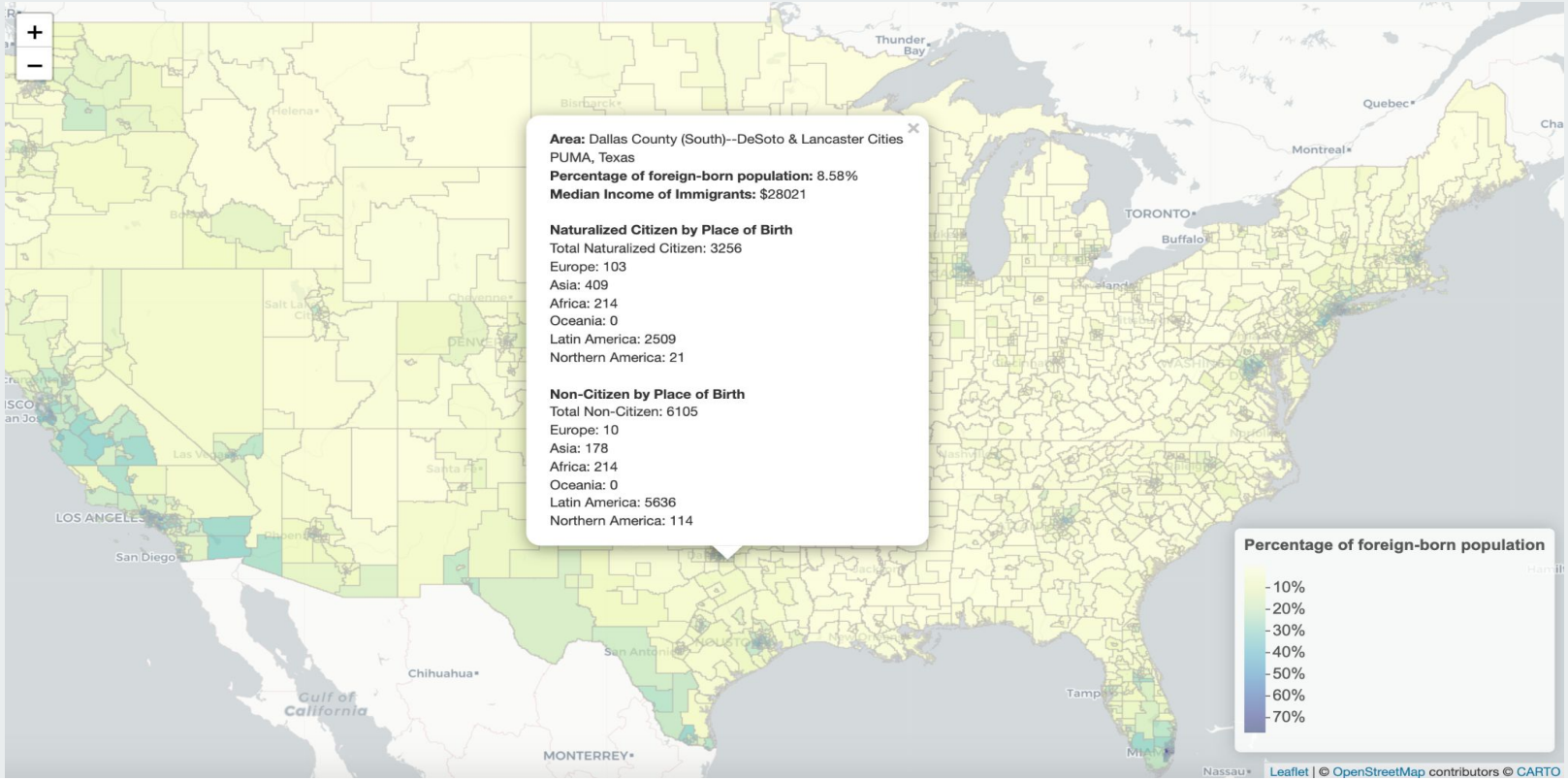


Top 10 Lowest



Interactive Map

We created an interactive map to visualize the percentage of the immigrant population and basic information of immigrants in each Public Use Microdata Area (PUMA) using 2015-2019 5-year American Community Survey estimates.



Future Work

- Explore the methodology for sharing the interactive map online
- Reach out to the immigrant organizations, such as Mayor's Office Immigrant Advancement, to see how can we put this interactive map to use
- Because of the time limit, our model only considers the difference in income level by state. For future work, we can convert city-level income data to PUMA and add a covariate for the median income by PUMA.

Reference

1. Chiswick, Barry R. “The Effect of Americanization on the Earnings of Foreign-Born Men.” Journal of Political Economy, vol. 86, no. 5, University of Chicago Press, 1978, pp. 897–921, http://www.jstor.org/stable/1828415.

2. U.S. Census Bureau, 2019 American Community Survey 1-Year Estimates, Public Use Microdata Sample (PUMS)

3. U.S. Census Bureau, 2015-2019 American Community Survey 5-Year Estimates, Public Use Microdata Sample (PUMS)

4. R survey package documentation, retrieved from https://cran.r-project.org/web/packages/survey/survey.pdf