

Methods

This project investigated both household and person-level characteristics by a sampled dataset from American Community Survey (ACS). At first, the original household-level dataset, which contained 2500 observations, was used to obtain the household-level demographic characteristics in terms of diverse variables. Data Dictionary was used to find the variable name in the dataset which could be processed in R. After computing the frequencies and percentages for each value of the variable, the following results analysis mapped the value back to the label in plain language. The missing values were assigned as “NA” in some cases to calculate the frequency and percentage.

To obtain Person-Level Demographic Characteristics, a new dataset *mydata* was created by joining the household-level and personal-level data tables by the same serial number (“*SERIALNO*”). The new dataset *mydata* had 6042 observations and 497 variables. Variables (columns) were selected to compute the frequencies and percentages and to calculate the means and standard deviations. The missing values of each variable were removed during calculating mean and standard deviation but were assigned as “NA” during calculating the frequency and percentage. Noted that “Weeks worked within the past year” was a categorical variable according to Data Dictionary, so each categorical response had to be converted to be numeric by assuming the midpoint of interval. Furthermore, average hourly wages are computed by $\frac{WAGP}{WKHP \times WKW}$, where *WAGP*, *WKHP* and *WKW* stand for salary income past 12 months, Usual hours worked per week, Weeks worked during past 12 months, respectively.

To obtain the Characteristics by Survey Response Mode among Single Persons, the *single_person* dataset was built by joining the household with number of person (*NP*) equals 1 and the whole Person-Level dataset by the same “*SERIALNO*”, resulting in 836 observations. The *single_person* was then split into three parts by the response modes (*single_model-3*), which were then used to analyze the single person characters.

Husband and wife dataset was built from *mydata* by selecting *HHT* (Household Family Type) equal to 1 and *RELP* (Relationship) equal to 0 or 1. The dataset also excluded the same-sex couples by selecting the household with a male husband and a female wife (*SEX*), resulting in 1183 couples to further analysis.

In order to obtain the relation between the couple average age and salary gap, husband and wife datasets were joined as a wide table by the same “*SERIALNO*”. The average age of each couple was computed by averaging the wife and husband’s age. The hourly wage gap was

obtained by equation $\left[\frac{WAGP}{(WKHP * WKW)} \right]_{husband} - \left[\frac{WAGP}{(WKHP * WKW)} \right]_{wife}$. After the plot and regression line were generated, the relation of them could be observed by the plot, equation of regression line and R-square. Note that the couple who both do not have any income were excluded in order to see more accurate relation.

Results

Table 1 showed the household-level demographic characteristics for selected variables. There were three regions, which were “Northeast, south and west” shown in the household-level data. Among those regions, most households were from West and least household were from Northeast. There were 0-10 bedrooms among different households. About half of households had 3 bedrooms and 2 bedrooms. For units in structure, most families (66.04%) lived in one family household. Most families were married couple type (47.4%). Most families had no children (64.04%) and 2 persons (29.04%) in family.

Table 2 showed the person-level demographic characteristics for selected variables. First, sex had almost the same distribution (49.02% male vs. 50.98% female). For the marital status, 42.63% of persons got married, which was almost equal to the percentage of never married (41.97%). Most of persons never served the military. For relationship, married couple and biological daughters and sons counted for 81.69% of persons. Educational attainment was evenly focused from GED to bachelor degree. The wages in the past 12 months had the standard deviation of 46432.2, indicating that the wages among the sample are highly imbalanced.

Table 3 showed the characteristics for the survey response mode among single persons. For the age variable, younger people preferred Internet (with age mean of 53.41), and Mail was used by older group (with age mean of 62.4). 61.72% of Mail single respondents are female, and the sex difference among Internet respondents was not significant. Employee from private company, business, or organization had the largest percentage among each response mode (39.32%, 49.08%, 48.78% for Mail, CATI/CAPI, and Internet, respectively).

Table 4 showed the husband and wife employment characteristics among married couples. For the wage variable, husbands had the wages with the mean of \$44614.6, nearly double of wives (with mean of \$23539.22). Both had extremely high standard deviation, meaning the income among both husband and wife were highly imbalanced. Husbands and wives had the similar commute time and husbands worked about 7 hours more than wives per week. In terms of worker class, husbands in government have the highest average hourly pay than any other classes for both husbands and wives. The science and engineering degree showed significant

difference among husbands and wives: 51.47% husbands have such degree but only 26.21% for wives, which indicated male might preferred to learn science and engineer than female do.

Figure 1 showed the relation between the couple average age and the difference in salary. The fitted linear regression line had been showed as $y=0.0042x+45.5801$, which had a very small slope. In addition, the R-squared was 0.000374, which was close to 0. Therefore, result showed that there was no statistically significant correlation between the couple average age and the difference in salary.

Table 1. Household-Level Demographic Characteristics

- Region Code*

	Northeast	South	West	NA
Frequency	95	1074	1331	0
Percentage	3.80%	42.96%	53.24%	0

- Number of bedrooms*

	0	1	2	3	4	5
Frequency	49	217	604	943	374	73
Percentage	1.96%	8.68%	24.16%	37.72%	14.96%	2.92%
	6	7	8	9	10	NA
Frequency	15	10	2	2	4	207
Percentage	0.60%	0.40%	00.08%	0.08%	0.16%	8.28%

- Units in structure:*

	One family house	Apartment	Other	NA
Frequency	1651	486	156	207
Percentage	66.04%	19.44%	6.24%	8.28%

- Household/family type*

	Married couple	Other:Male	Other:Female	Nonfamily: Male alone
Frequency	1185	98	247	260
Percentage	47.40%	3.92%	9.88%	10.40%
	Nonfamily: Male Not alone	Nonfamily: Female alone	Nonfamily: Female Not alone	NA
Frequency	78	369	56	207
Percentage	3.12%	14.76%	2.24%	8.28%

- Presence and age of children*

	< 6 years	6 ~ 17 years	< 6 & 6 ~ 17 years	No children	NA
Frequency	157	387	148	1601	207
Percentage	6.28%	15.48%	5.92%	64.04%	8.28%

- When moved into house or apartment*

	< 10 years ago	10 or more years ago	NA
Frequency	1277	1016	207
Percentage	51.08%	40.64%	8.28%

- Number of persons in family*

	2 persons	3 persons	4 persons	5+ persons	NA
Frequency	726	290	297	217	970
Percentage	29.04%	11.60%	11.88%	8.68%	38.80%

- Family income past 12 months (consider any losses to be equal to zero income)*

	< \$30,000	\$30,000 - \$49,999	\$50,000 - \$69,999	70,000 - \$99,999	\$100,000 or more	NA
Frequency	321	255	239	254	461	970
Percentage	12.84%	10.20%	9.56%	10.16%	18.44%	38.80%

- Type of unit*

	Housing unit	Institutional group quarters	Non-institutional group quarters	NA
Frequency	2293	108	99	0
Percentage	91.72%	4.32%	3.96%	0

- Response mode*

	Mail	CATI/CAPI	Internet	NA
Frequency	1230	675	388	207
Percentage	49.20%	27.00%	15.52%	8.28%

Table 2. Person-Level Demographic Characteristics

- Sex*

Sex	Male	Female
Frequency	2962	3080
Percentage	49.02%	50.98%

- Marital status*

Marital status	Married	Widowed	Divorced	Separated	Never married
Frequency	2576	337	508	85	2536
Percentage	42.63%	5.58%	8.41%	1.41%	41.97%

- Class of worker*

Class of worker	Employee of a private company, business or organization	Government employee	Self-employed	Working without pay or unemployed	NA
Frequency	2432	577	398	60	2575
Percentage	40.25%	9.55%	6.59%	0.99%	42.62%

- Military Service*

Military Service	Now on active duty	On active duty in the past	Only on active duty for training in Reserves
Frequency	28	498	51
Percentage	0.46%	8.24%	0.84%
Military Service	Never served in the military	less than 17 years old (NA)	
Frequency	4233	1232	
Percentage	70.06%	20.39%	

- Relationship (F: frequency; P: percentage)**

	Reference person	Husband/wife	Biological son or daughter	Adopted son or daughter	Stepson or stepdaughter
F	2293	1185	1458	40	92
P	37.95%	19.61%	24.13%	0.66%	1.52%
	Brother or sister	Father or mother	Grandchild	Parent-in-law	Son-in-law or daughter-in-law
F	68	72	165	26	25
P	1.13%	1.19%	2.73%	0.43%	0.41%
	Other relative	Roomer or boarder	Housemate or roommate	Unmarried partner	Foster child
F	109	27	81	127	5
P	1.80%	0.45%	1.34%	2.10%	0.08%
	Other nonrelative	Institutionalized group quarters population	Noninstitutionalized group quarters population		
F	62	108	99		
P	1.03%	1.79%	1.64%		

- Educational attainment**

<i>Education</i>	No schooling completed	Preschool & Kindergarten	Elementary school
Frequency	217	155	570
Percentage	3.59%	2.57%	9.48%
<i>Education</i>	Middle & High school	High school diploma & GED	College students (no degree)
Frequency	848	1213	1143
Percentage	14.04%	20.08%	18.93%
	Associate's & Bachelor's degree	Graduate students & Professional degree	Less than 3 years old (NA)
Frequency	1149	543	204
Percentage	19.02%	8.99%	3.38%

- *Age in years ; Wages or salary income in the past 12 months (use WAGP) ; Usual hours worked per week past 12 months (use WKHP) ; Weeks worked within the past year (use WKW) ; Average hourly wages or salary income throughout the past 12 months during weeks worked*

Variable	Mean	Standard Deviation
<i>Age in years</i>	40.0134	23.4079
Wages or salary income in the past 12 months	24554.57	46432.2
Usual hours worked per week past 12 months	38.029	21.0629
Weeks worked within the past year	44.6073	13.1601
Average hourly wages or salary income throughout the past 12 months during weeks worked	23.0246	32.3273

Table 3. Characteristics by Survey Response Mode among Single Persons

- Mean, standard deviation, median, first quartile, and third quartile for age in years

Response Mode	Mean	Standard Deviation	Median	1st Quartile	3rd Quartile
Mail	62.4	17.46645	64	52	76
CATI/CAPI	57.55	19.06556	58	44	71
Internet	53.41	18.18285	54	38	67

- Frequencies and percentages for sex

Response mode	Sex	Frequency	Percentage
Mail	Male	147	38.28%
	Female	237	61.72%
CATI/CAPI	Male	71	43.56%
	Female	92	56.44%
Internet	Male	42	51.22%
	Female	40	48.78%

- Frequencies and percentages for class of worker
 - Response mode: Mail (contains: NA 161 (0.4193))

Class of Worker	Frequency	Percentage
Employee of a private company, business or organization	151	39.32%
Government employee	44	11.46%
Self-employed	27	7.03%
Working without pay or unemployed	1	0.26%

- Response mode: CATI/CAPI (contains NA 52 (0.3190))

Class of Worker	Frequency	Percentage
------------------------	------------------	-------------------

Employee of a private company, business or organization	80	49.08%
Government employee	20	12.27%
Self-employed	11	6.75%
Working without pay or unemployed	0	0

- Response mode: Internet (contains NA 18 (0.2195))

Class of Worker	Frequency	Percentage
Employee of a private company, business or organization	40	48.78%
Government employee	17	20.73%
Self-employed	7	8.54%
Working without pay or unemployed	0	0

Table 4. Husband and Wife Employment Characteristics among Married Couples

- Wages or salary income in the past 12 months

Husband/Wife status	Mean	Standard Deviation	# of Record
Husband	44614.6	68247.99	1183
Wife	23539.22	41438.39	1183

- Travel time to work

Husband/Wife status	Mean (min)	Standard Deviation (min)	# of Record
Husband	28.25	23.21	1183
Wife	25.61	21.79	1183

- Usual hours worked per week past 12 months

Husband/Wife status	Mean (min)	Standard Deviation (min)	# of Record
Husband	43.01	12.05	1183
Wife	36.24	12.32	1183

- Average hourly wages or salary income throughout the past 12 months

Husband/Wife	Class of Worker	Mean	Standard Deviation	# of Record
Husband	Employee of a private company, business or organization	32.42	30.61	626
	Government employee	36.32	48.23	144
	Self-employed	18.71	35.77	153
	Working without pay or unemployed	0	NA	10
Wife	Employee of a private company,	24.85	23.89	531

	business or organization			
	Government employee	24.46	21.18	175
	Self-employed	22.99	84.67	103
	Working without pay or unemployed	19.61	27.73	7

- Field of degree science and engineering (Not include NA)

Husband/Wife	Science and engineering	Frequency	Percentage
Husband (1183 obs)	YES	228	51.47%
	NO	215	48.53%
Wife (1183 obs)	YES	108	26.21%
	NO	304	73.79%

Figure 1. Plot of Couple Average Age by Difference in Salary

