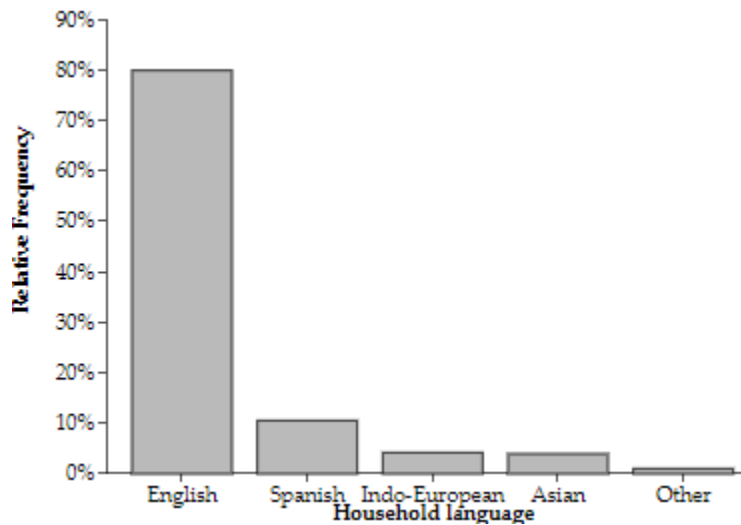


1. There are 18 variables recorded for each household.

Categorical: Division, region, access to the internet, whether they are a food stamp/SNAP recipient, type of household fuel used, whether they have a computer, whether they have a smartphone, number of vehicles (treated as categorical with numeric codes, such as 6 = 6 or more vehicles), year their home was built, household language, how long ago they moved into their house/apartment (treated as categorical with numeric codes, such as 1 = 12 months or less), response mode.

Quantitative: Number of people in the household, number of bedrooms, monthly rent (\$), property value (\$), household income over the past 12 months, number of biological children in household.

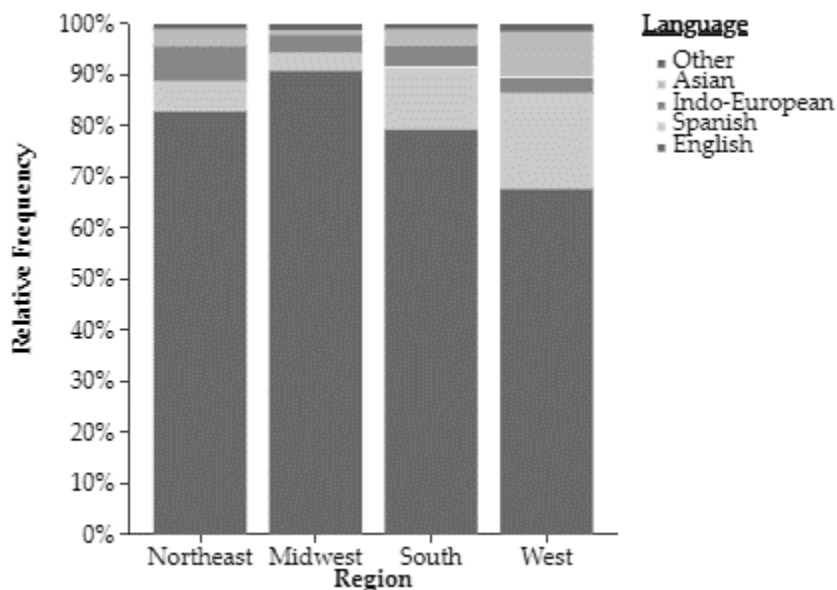
2. The bar graph reveals that the vast majority of the surveyed households speak English (80.03%) and the second most commonly spoken language is Spanish (10.53%). Other Indo-European languages and Asian and Pacific Island languages are spoken with about the same relative frequency (About 4%). Other languages account for about 1% of households surveyed.



| Category Name | Frequency | Relative Frequency |
|---------------|-----------|--------------------|
| English | 2401 | 80.03% |
| Spanish | 316 | 10.53% |
| Indo-European | 129 | 4.3% |
| Asian | 121 | 4.03% |
| Other | 33 | 1.1% |
| Total | 3000 | 100% |

3. The two-way table provides a numerical summary of the frequency and relative frequency of spoken language for households in each region.

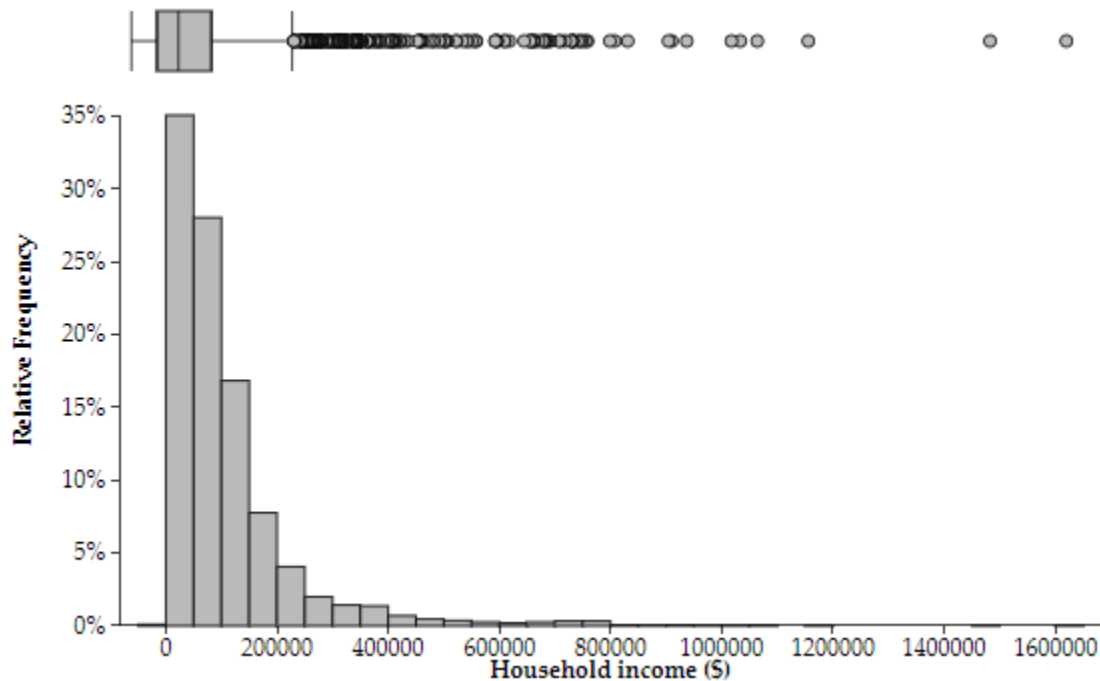
| | | Region | | | | |
|----------|---------------|------------|------------|-------------|-------------|-------------|
| | | Northeast | Midwest | South | West | Total |
| Language | English | 440 (83%) | 603 (91%) | 904 (79.4%) | 454 (67.9%) | 2401 (80%) |
| | Spanish | 30 (5.7%) | 23 (3.5%) | 138 (12.1%) | 125 (18.7%) | 316 (10.5%) |
| | Indo-European | 38 (7.2%) | 23 (3.5%) | 48 (4.2%) | 20 (3%) | 129 (4.3%) |
| | Asian | 18 (3.4%) | 6 (0.9%) | 37 (3.3%) | 60 (9%) | 121 (4%) |
| | Other | 4 (0.8%) | 8 (1.2%) | 11 (1%) | 10 (1.5%) | 33 (1.1%) |
| | Total | 530 (100%) | 663 (100%) | 1138 (100%) | 669 (100%) | 3000 (100%) |



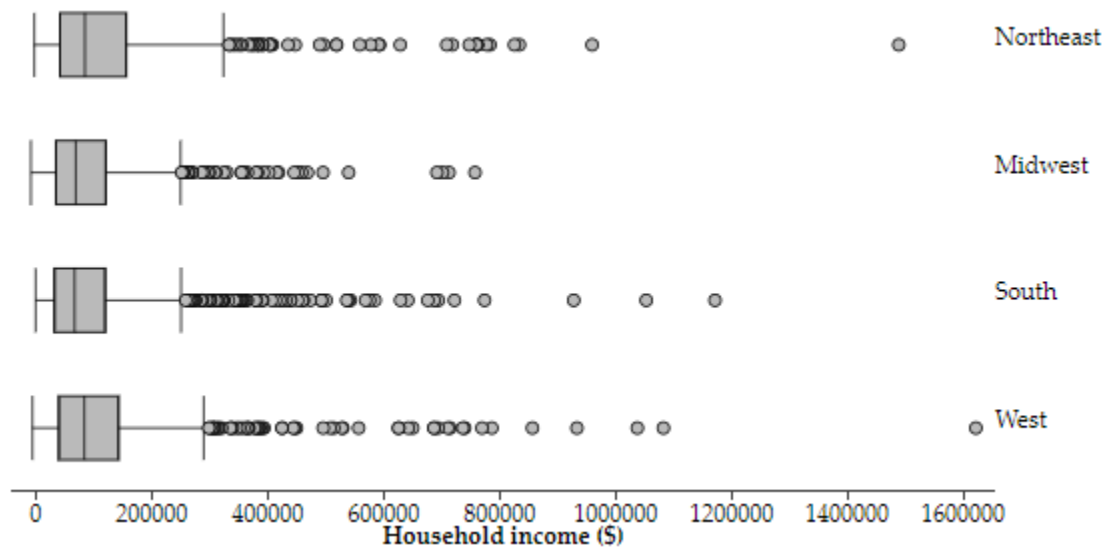
The segmented bar graph reveals that there is an association between region and language spoken. Although households from all 4 regions most commonly spoke English, the percentage of English-speaking households was greatest for households from the Midwest (91%), followed by the Northeast (83%), then the South (79.4%) and West (67.9%). The percentage of Spanish-speaking households was greatest for those in the West (18.7%), followed by the South (12.1%), the Northeast (5.7%), and the Midwest (3.5%). In the Midwest, other Indo-European language speaking households is about the same as the Spanish speaking households (3.5%), and in the West, Asian and Pacific Island language speaking households were more common (9%) than Other Indo-European language speaking households (3%).

4. The distribution of household income among ACS survey respondents is strongly skewed to the right with a single peak in the \$0 to < \$50,000 income interval. There are many high outliers, as shown in the boxplot. These high outliers consist of 205 households who earned over \$273,762.50 in the past year. There were 3000 households surveyed, which means that $\frac{205}{3000} \approx 0.0683$ or 6.83% of households surveyed have annual incomes that are high outliers according to the $1.5 \times IQR$ rule. The maximum household income reported was \$1,622,000. Also noteworthy is that three households reported negative household incomes: -\$8200, -\$5500, and -\$2300. These negative household incomes were not low outliers.

The mean household income was \$108,509.14 and the median household income was \$73,850. The standard deviation of household incomes was \$128,462.94 and the interquartile range was $Q_3 - Q_1 = \$131,225 - \$36,200 = \$95,025$. So, the range of the middle 50% of household incomes was \$95,025.



5. Here are parallel boxplots to display the distribution of household income for the 4 regions of the United States, followed by summary statistics.



| Group Name | <i>n</i> | mean | SD | min | <i>Q</i> ₁ | med | <i>Q</i> ₃ | max |
|--------------|----------|------------|------------|-------|-----------------------|--------|-----------------------|-----------|
| 1: Northeast | 530 | 131,228.13 | 157,044.33 | -2300 | 42,000 | 85,000 | 156,200 | 1,489,000 |
| 2: Midwest | 663 | 95,237.83 | 95,323.41 | -8200 | 35,000 | 69,200 | 121,000 | 758,000 |
| 3: South | 1138 | 97,039.14 | 112,021.69 | 0 | 32,400 | 67,200 | 120,000 | 1,172,000 |
| 4: West | 669 | 123,173.81 | 152,217.76 | -5500 | 40,000 | 84,000 | 143,000 | 1,622,000 |

For all four regions, the distribution of household income is skewed to the right with many high outliers. The median household income is greatest for the Northeast (\$85,000), followed by the West (\$84,000), the Midwest (\$69,200), and finally the South (\$67,200). The household incomes vary the most for households in the Northeast ($IQR = \$114,200$), followed by the West ($IQR = \$103,000$), and are similar for the South ($IQR = \$87,600$) and Midwest ($IQR = \$86,000$). The distribution of household income is similar for households in the Northeast and West and is similar for households in the South and Midwest, but the distribution of household income in the Northeast and West differ from the South and Midwest.

There is also a large disparity between the regions among the households with the greatest incomes. The maximum household income in the Midwest was only \$758,000, which is less than half of the highest household income value for the west (\$1,622,000) and almost half of the highest household income value in the Northeast (\$1,489,000). There were several households in the Northeast, South, and West that earned more than the maximum household income in the Midwest.

Extension: Answers will vary.