

PS1_Part1

Nora Nickels

4/14/2018

Problem Set 1

Part 1: Write a data section for your assigned data set

Data

The data used in this study, data from the Current Population Survey (CPS), was accessed directly from the National Bureau of Economic Research website [http://nber.org/data/cps_basic.html]. It can be accessed there, under the Data tab. The original source comes from the U.S. Bureau of Labor Statistics (BLS), which jointly sponsors the CPS with the U.S. Census Bureau. The BLS maintains a CPS home page [<https://www.census.gov/programs-surveys/cps.html>], where the data can be accessed as well. It is maintained and curated there, and can be accessed in many forms. A list of variable information can be found here: http://nber.org/cps-basic/January_2017_Record_Layout.txt. For the purpose of this study, the June 2017 csv file of the CPS data was used; many formats can be accessed here: http://thedataweb.rm.census.gov/ftp/cps_ftp.html.

The CPS database has been utilized by several key papers, ranging in topic from health, income equality, poverty rates, and same-sex marriage (Burkhauser et al., 2011; Davis et al., 2015; Edwards & Lindstrom, 2017).

The CPS is conducted by collecting data on a monthly basis using a probability sample of about 60,000 occupied household. It is administered as a completely computerized document that is administered by field representatives, through personal and telephone interviews.

Table 1 (Figure 1) includes descriptives statistics highlighting eight key features of the CPS data set. These features include overall family income, highest level of school completed or degree received, race, number of own children < 18 years of age, total number of people living in the household, household type, whether or not the respondent owns a business/farm, and region.

	Family Income	Education Level	Race	# Of Children	# In Household	Household Type	Business/Farm Owner	Region
Median	8	11	23	11	3	6	4	3
Mean	8.28	8.624	17.335	10.0357	2.8585	6.2111	3.6633	2.7019
SE.mean	0.014	0.016	0.023	0.0045	0.0049	0.0066	0.003	0.0028
CI.Mean.0.95	0.027	0.031	0.045	0.0088	0.0095	0.013	0.0059	0.0054
Var	27.433	35.86	79.18	3.0175	3.5094	6.5342	1.328	1.1289
Std.dev	5.238	5.988	8.898	1.7371	1.8734	2.5562	1.1524	1.0625
Coef.var	0.633	0.694	0.513	0.1731	0.6554	0.4116	0.3146	0.3932

Figure 1: Table 1

Focusing on region and business / farm ownership, Figure 2 shows the distribution of samples from four regions of the US - the midwest, south, northeast and west. Within that, Figure 1 shows the distribution amongst those regions of individuals who either do not own or own either a farm or a business. We can see that most respondents in June of 2017 were from the south, and the least respondents were from the northwest. Further, the overwhelming majority of respondents do not own either a farm or a business, in all four regions.

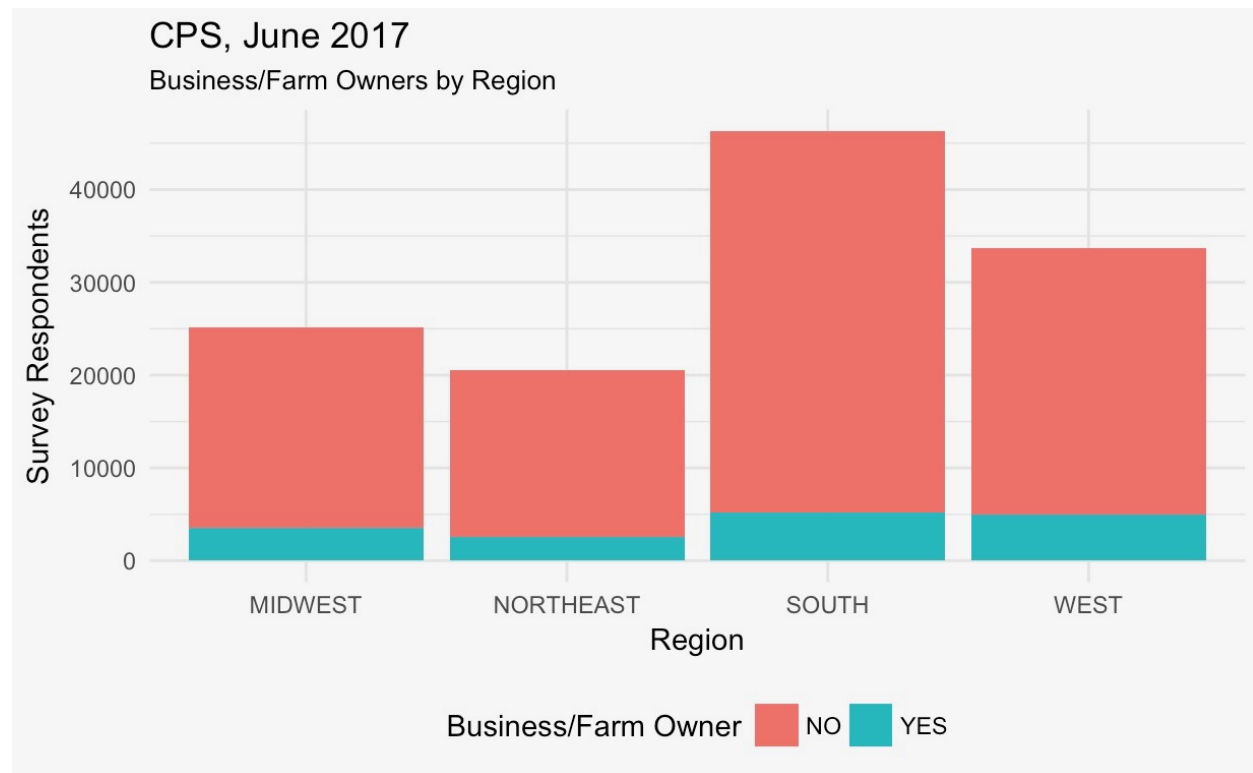


Figure 2: Region / Business and Farm Ownership

Figure 3 shows the distribution of household sizes amongst survey respondents in June of 2017. The majority of households contain three individuals, and household sizes dramatically taper off after 5 or 6 total members.

Show at least one conditional (slice) description of the data (e.g., all variable descriptive statistics by nationality of survey respondent). This can be a table or visualization.

References

- Burkhauser, R. V., Feng, S., Jenkins, S. P., & Larrimore, J. (2011). Estimating trends in US income inequality using the Current Population Survey: the importance of controlling for censoring. *The Journal of Economic Inequality*, 9(3), 393-415.
- Davis, W. W., Hartman, A. M., & Gibson, J. T. (2015). Trends in smoking prevalence by race based on the Tobacco Use Supplement to the Current Population Survey.
- Edwards, A., & Lindstrom, R. (2017). Measuring the presence and impact of same-sex married couples on poverty rates in the Current Population Survey (No. 2017-01). SEHSD Working Paper.

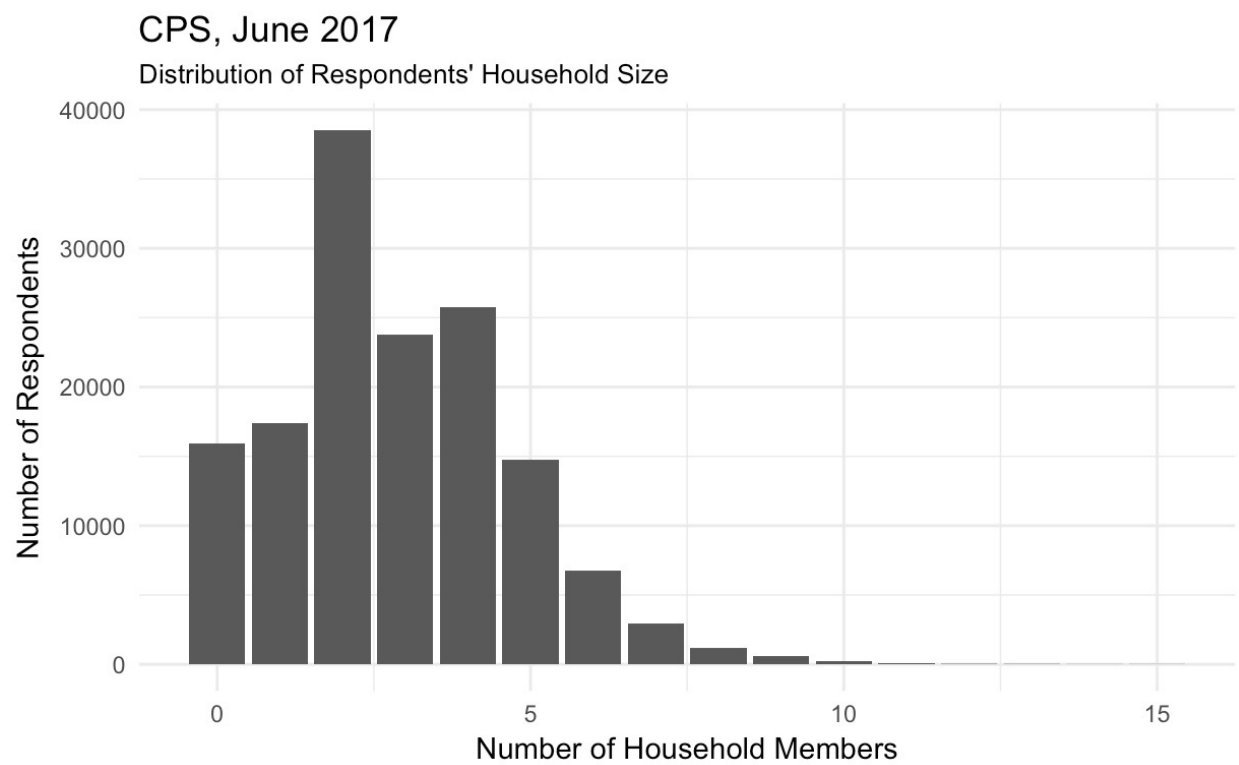


Figure 3: Distribution of Household Sizes