Problem Set 1
Weiwei Zheng
April 15, 2018

Part I

1. The Current Population Survey (CPS), sponsored jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS), is the primary source of labor force statistics for the population of the United States. The survey includes a representative sample of about 60,000 homes and focuses on those individuals who are 15 years and older to make an inferential assumption about the US population.

In addition to being the primary source of monthly labor force statistics, the CPS is used to collect data for a variety of other studies that keep the nation informed of the economics and social well-being of its people. The data is available on some well-known social science dataset websites including the United States Census Bureau, National Bureau of Economic Research, US Census Data for Social, Economic and Health Research, etc.

2. The CPS data has been studied by different social scientists since long. Arenas-Arroyo (2018) leveraged the data to study impacts of granting access to driving licenses to undocumented immigrants. Rinz and Voorheis (2018) used administrative earning data from the dataset and looked into the distributional effects of minimum wages. Gelber and co-authors also used it to examine the role of social security benefits in the initial increase of older women's employment.

3. Each month during interview week, field representatives and computer assisted telephone interviewers contacts and interviews a responsible person living in each sample unit selected to complete an interview. Typically, the week containing the 19th of the month is the interview week. In CPS, households are in sample for eight months. Each month, one-eighth of the households are in sample for the first time, one-eighth for the second time, etc. A personal visit interview is required for all first month-in-sample households because the CPS sample is strictly a sample of addresses.

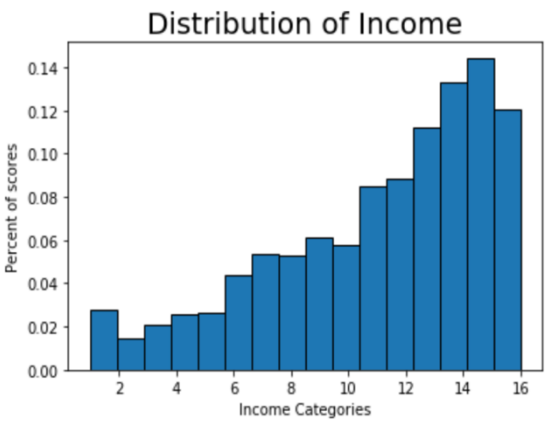
4. The following table provides a summary of descriptive statistics of eight most important variables.

Table 1. Summary Statistics

	Sex	**Education**	**Age**	**Income**	**Years in College**	**Work hours weekly**	**Number of Children**	**Race**
Count	131759	107189	131759	131759	28720	62371	131759	131759
Mean	Female (mode)	High school (mode)	39	100-150 thousand (mode)	2.8	36.1	0.38	White only (mode)
Std	NA	NA	23.1	NA	0.99	15.1	0.87	NA
Min	NA	NA	15	NA	1	0	0	NA
Max	NA	NA	85	NA	5	99	10	NA

5. The codebook for different income categories are as such:

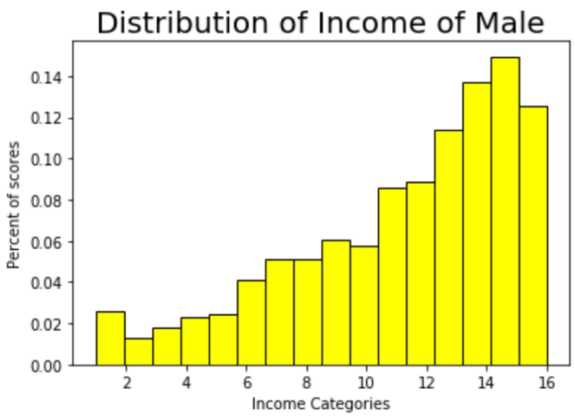
1	LESS THAN $5,000
2	5,000 TO 7,499
3	7,500 TO 9,999
4	10,000 TO 12,499
5	12,500 TO 14,999
6	15,000 TO 19,999
7	20,000 TO 24,999
8	25,000 TO 29,999
9	30,000 TO 34,999
10	35,000 TO 39,999
11	40,000 TO 49,999
12	50,000 TO 59,999
13	60,000 TO 74,999
14	75,000 TO 99,999
15	100,000 TO 149,999
16	150,000 OR MORE

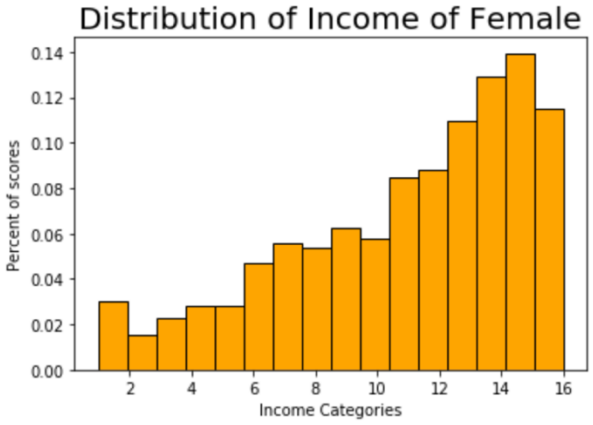


We can tell from the histogram that the mode is between 100 thousand to 150 thousand dollars a year and the mean seems to lie in somewhere around 60 to 80 thousand, which is extremely high. The graph perfectly displays the material prosperity of United States of America.

6. Income Distribution by Gender

We cannot see severe discrepancy between different genders on family income. And females seem to have higher average income but maybe fewer super rich people.





Citations

Arenas-Arroyo, Esther. "Labor Market Impacts of Granting Access to Driving Licenses to Undocumented Immigrants." (2018).

Gelber, A., Isen, A., & Song, J. (2018). The Role of Social Security Benefits in the Initial Increase of Older Women's Employment. *Women Working Longer: Increased Employment at Older Ages*, 239.

Rinz, Kevin, and John Voorheis. "The Distributional Effects of Minimum Wages: Evidence from Linked Survey and Administrative Data." (2018).