Task 1 from predoc.org

Labor Force Participation from 1976 to 2015 Data Source: U.S. Current Population Survey (CPS)

Question 1 (a) Please summarize the key trends for wages and labor force participation.

The "lfp_timeseries" plot examines aggregate labor force participation between 1977 and 2015. There is a significant jump in individuals observed both in the labor force and not in the labor force around 2001. This is consistent with the 2001 survey sample increase of 10,000 (https://www.bls.gov/cps/cpsjul2001.pdf). The data shows some variation over time consistent with sample size changes. Dips in observations in the labor force are consistent with dips not in the labor force and vice versa up to the 2000 mark. Post-2000 we see a dip in labor force participation and a rise in observations not in the labor force. This may be associated with economic downturns, in which individuals left the labor market entirely. Additionally, the data may reflect a change in population age breakdowns. As the population ages, without equal replacement, the number of individuals not in the labot market would increase and the number of individuals in the labor market would decrease. This is consistent with observations from the Bureau of Labor Statistics (https://www.bls.gov/opub/mlr/2016/article/labor-force-participation-what-has-happened-since-the-peak.htm).

The "wages_timeseries" plots total wages over time from 1977 to 2015. The plot indicates relatively constant growth in mean and median wages over time, with a slight divergence in mean and median starting around 2000. The mean wage appears to increase at a higher rate than the median wage. This is constistent with an increasingly right-skewed distribution of wages. The increasing difference in mean and median wages over time may be associated with changes in high and low income brackets, where high income individuals see greater wage growth than middle and low income individuals.

(b) Among men older than age 25, which groups of people have had the biggest changes in labor force participation?

The "not_lf_state" and "lf_state" plots show mean year-to-year percent changes in labor force participation for men over 25 by the five states with the smallest mean percent change and the five states with the largest mean percent change. The District of Columbia saw the greatest mean percent change in observations in the labor force and not in the labor force. However, this could be attributed to a change in the number of observations in DC, either from a population increase or a survey sampling change. Three states - Michigan, Ohio, and Pennsylvania - saw an average year-to-year decrease in labor force participation. However, Ohio and Pennsylvania were also among the states

with smallest mean changes in individuals not in the labor force. Hawaii and Delaware were among the top five states for increases in people not in the labor force and in the labor force. Again, this may be attributed to population shifts.

Broken down into age groups, men in the 65+ category saw the smallest change in labor force participation, with largely consistent numbers in the labor force and not in the labor force, as seen in the "lfp_age_group" plot. For all three age group categories (25-45, 45-65, 65+), the number of observations of not in the labor force is fairly stable over time. The 25-45 and 45-65 categories indicate a jump in the labor force around 2001, consistent with the change in survey sampling. In the post 2001 era, there is a downtun in the number of 25-45 yearo olds in the labor force and an upturn in the number of 45-65 year olds in the labor force. There is no observable, opposite reaction in the "not in the labor force" category, indicating an aging population.

(c) What factors do you think are driving these patterns? What evidence might you want to assemble to test these hypotheses if you were to investigate them further?

As noted above, many of the distinct changes in observed labor force participation can be attributed to population shifts and changes in survey sampling. Additionally, changes in survey models make it difficult to compare changes among racial and education groups, as data collection and variable categories changed over time. With additional time and information, it would be valuable to recategorize and standardize racial groups ("race"), education levels ("educ"), and employment status ("empstat").