

SOCI 385 - Social Statistics

Fall 2023

Assignment 5

This assignment is due via Canvas by 10:00 AM on Tuesday, November 7, 2023. Please show all your code and write your responses in complete sentences as appropriate. Submit your rendered pdf and your Quarto notebook (the .qmd file). You should use the R shortcuts for the hypothesis tests.

For this assignment, you will explore some of the data and research questions that sociology major Hanwen Zhang is using for his senior project. The dataset is available **here** and comes from the 2021 GSS.

The variables are:

- **born**: whether the respondent is born in the US (labeled **Native**) or born in a different country (labeled **Immigrant**);
- **age**: respondent's age in years, from 18-89;
- **yearsusa**: respondent's number of years in the US, from 1-83, only available for immigrants;
- **ophrdwrk_essential**: binary variable noting if respondent said that hard work is essential for getting ahead in life, coded 1 if yes and 0 if they chose any of the other possible answers (hard work is very important, fairly important, not very important, or not important at all);
- **opknow_essential**: binary variable noting if respondent said that knowing the right people is essential for getting ahead in life, coded 1 if yes and 0 if they chose any of the other possible answers (knowing the right people is very important, fairly important, not very important, or not important at all);
- The following **op*** variables are other questions asking respondents how important each is for getting ahead in life. Each has the possible responses of: "Essential," "Very Important," "Fairly Important," "Not Very Important," or "Not Important At All." The variables are:

- **opwlth**: need a wealthy family to get ahead;
- **oppared**: need educated parents to get ahead;
- **opeduc**: need good education to get ahead;
- **opclout**: need political connections to get ahead;
- **oprace**: need to be the right race to get ahead;
- **oprelig**: need to be the right religion to get ahead;
- **opsex**: need to be the right sex to get ahead.

Questions

1. Test if there is a significant difference at the 95% confidence level in the proportion of immigrants vs native born respondents who say hard work is essential. Interpret your results in a sentence.
2. Test if there is a significant difference at the 95% confidence level in the average age between immigrants and native born respondents. Interpret your results in a sentence.
3. Consider immigrants only, and test if there is a significant difference at the 95% confidence level in average years in the US between immigrants who say hard work is essential and immigrants who do not say hard work is essential. Interpret your results in a sentence.
4. Consider immigrants only, and create a binary variable distinguishing those who have been in the US less than fifteen years from those who have been in the US fifteen years or more. Test if there is a significant difference between these groups at the 95% confidence level in the proportion saying knowing the right people is essential for getting ahead. Interpret your results in a sentence.
5. From the following variables, choose one where you would hypothesize immigrants and native born respondents would have significantly different responses: `opwlth`, `oppared`, `opeduc`, `opclout`, `oprace`, `oprelig`, or `opsex`. In a sentence, describe why you have this hypothesis. From your chosen variable, create one new binary variable distinguishing responses of “Essential” or “Very Important” from all other responses. Then use this new variable to test if there is a significant difference at the 99% confidence level in the proportion responding “Essential” or “Very Important” (rather than a different response) between immigrants and native born respondents. Interpret your results in a sentence.