Instructions for Reproducing Analysis

Mohsen Monji

2024-09-16

Introduction

This document provides an overview of the steps required to reproduce the analysis of disparities in self-rated mental health among young adults (aged 18-39) in Canada using the 2017 General Social Survey-Public Microdata Files. The analysis is organized into three stages using three R scripts:

1. `Data Wrangling`: Cleaning and recoding variables.
2. `Descriptive Statistics`: Producing frequencies and proportions of self-rated mental health by various socio-demographic variables among young adults (aged 18-39).
3. `Chi-Square Tests and Logistic Regression`: Analyzing associations between self-rated mental health and predictors using chi-square tests and logistic regression.

Step 1: Download Files.zip in this repository

Download the Files.zip in this repository, unzip it and save it in a directory of your choice on your machine.

Step 2: Download the GSS 2017 Dataset

The GSS-2017 dataset is not included in the repository due to licensing restrictions. Download it from ODESI at [https://odesi.ca/] and follow the steps below:

1. Unzip the file and rename it to `gss2017.csv`.
2. Move the `gss2017.csv` file to the main directory that you placed the R scripts in the Files.zip.

Step 3: Run R Scripts

Follow the steps below to reproduce the analysis:

3.1 Data Wrangling

First, run the `Code for Data Wrangling.R` script. This will recode the variables of interest, drop cases with missing values, and output a cleaned dataset with only the selected variables named (`gss_2017_selected.csv`), which is used in the subsequent analysis. The `gss_2017_selected.csv` will be saved in the main directory where the R scripts are located.

3.2 Descriptive Statistics

Second, run the `Code for Descriptive Statistics.R` script. It will first filter the saved dataset from the previous step to include only young adults (aged 18-39) and then will create a sample characteritics table(table 1) and calculate the frequencies and percentages of self-rated mental health by all predictors such as sex, marital status, household income, etc(table 2).

3.3 Chi-Square Tests and Logistic Regression

Third, run the `Code for Logistic Regression.R` script. It will perform chi-square tests for associations between self-rated mental health and the categorical predictors. It will also fit a logistic regression model to further explore these associations adjusting for confounders. The script will display the results of the chi-square tests and the logistic regression model, and will create a table of the logistic regression output(table 3).

Conclusion

This analysis explored the associations between self-rated mental health and demographic and socioeconomic predictors among young adults aged 18-39 in Canada, using data from the 2017 General Social Survey, Public Microdata Files. The results show that **being female** (OR = 1.49, 95% CI: 1.23, 1.80, $p < 0.001$) and **being single** (OR = 1.86, 95% CI: 1.48, 2.35, $p < 0.001$) were significantly associated with higher odds of reporting poor or fair mental health among young adults. Similarly, young adults who were **divorced or widowed** had higher odds of poor mental health (OR = 1.98, 95% CI: 1.24, 3.06, $p = 0.003$). On the other hand, **higher household income** was protective against poor mental health, with those earning $100k or more having significantly lower odds (OR = 0.60, 95% CI: 0.47, 0.77, $p < 0.001$). **Being a landed immigrant** was also associated with lower odds of reporting poor/fair mental health (OR = 0.67, 95% CI: 0.46, 0.97, $p = 0.036$), while no significant associations were found for visible minority status, middle household income, or lower levels of educational attainment. The results also show that age is not a significant predictor of poor or fair mental health among young adults in Canada(OR = 0.99, 95% CI: 0.97, 1.01, $p = 0.197$).

Final Words

By following the steps outlined above, you will be able to reproduce the analysis of disparities in self-rated mental health among young adults (aged 18-39) in Canada using the 2017 General Social Survey data. Ensure the dataset is placed in the correct directory, and run the R scripts to reproduce this analysis.

If you encounter any issues, feel free to reach out for assistance!