

Replication Project Diary

This replication project is based on the article “Race/Ethnicity and Family Contact: Toward a Behavioral Measure of Familialism” by Joseph Comeau from 2012. I will be replicating Tables 1 and 2 of descriptive statistics and linear regression from the article.

2024-01-07

Created folder within elon folder on desktop -> jterm 2024 -> replication project. Stored **GSS_Codebook** and **Comeau-2012-Race&Familialism** under **resources** folder in **replication project** folder

2024-01-08

Created proposal document titled **arakes-repl01-proposal-2024-01-08**

Started identifying variables I need to replicate through the use of **GSS_Codebook** and the **Comeau-2012-Race&Familialism** article. Noted them in **arakes-variables-2024-01-08**.

2024-01-09

In **arakes-variables-2024-01-08** I wrote all variables from **GSS_Codebook** used in analysis in **Comeau-2012-Race&Familialism**

Created timeline on page 1 of **arakes-repl01-proposal-2024-01-08**.

Began writing the proposal document. Established that my mnemonic used for all of my files will be “repl”. Also established the title of data scripts as **arakes-repl-data1.R** and title of analysis script files as **arakes-repl-analysis1.R**.

2024-01-10

Finished **arakes-repl01-proposal-2024-01-08**. Identified how my 5 data management script files and my 3 analysis script files will be split up. Established steps that will need to be taken to clean data and steps taken to analyze data. Also established final paper format and organizational plans.

Version control will be done through the posted folder. Files backed up through one drive.

2024-01-11

Created **arakes-repl-data1.R**. Wrote code for first data management script. Subsetting larger dataset into just the variables I need. Saved new dataset as **arakes-repl-data2.RDS**

2024-01-15

Created **arakes-repl-data2.R**. Input new dataset **arakes-repl-data2.RDS** and began more data management. Created dummy variables for sex, marital status, race, and geographic mobility and saved as new dataset titled **arakes-repl-data3.RDS**

2024-01-16

Created **arakes-repl-data3.R** and input dataset from yesterday: **arakes-repl-data3.RDS**. Continued data management by recoding all the family variables. Changed scale numbers for each. Saved as a new dataset titled **arakes-repl-data4.RDS**.

Had trouble creating average of variables, had to change code on **arakes-repl-data3.R**. Switched from using a factor statement to recode my family variables to using a ifelse nested statement.

2024-01-17

Finished creating averages of nuclear family index and extended family index on **arakes-repl-data4.R** and **arakes-repl-data5.R**, respectively. Also changed a few variables to numeric in **arakes-repl-data5.R** so they can be used in my analysis. Saved final cleaned dataset as **arakes-repl-data6.RDS**

2024-01-18

Created first analysis script titled **arakes-analysis-data1.R**. Generated all descriptive statistics for the nuclear family variable, including means, standard deviations, and proportions. In **arakes-analysis-data2.R** I generated all descriptive statistics for extended family variables. In **arakes-analysis-data3.R** I generated all descriptive statistics for the individual family variables.

Created **arakes-analysis-data4.R** where I created regression models for each of the indexes and individual family variables.

2024-01-19

Created my two replicated tables in **arakes-repl-tables-hidden-2024-01-19** as well as the provenance shown version. Sheet 1 contains the descriptive statistics table and sheet 2 contains the table from the regression models.

2024-01-21

Wrote my final research paper titled **arakes-repl-paper-hidden-2024-01-22** as well as the version shown with provenance titled **arakes-repl-paper-shown-2024-01-22**.

2024-01-22

Created the read me file titled **arakes-repl-README-2024-01-22**.