Cooperative Election Study (CES) 2016-2024: Open-Ended Religion and Race Response Recoding

[Your Name]

1 Overview

This dataset contains recoded open-ended responses for religion and race questions from the Cooperative Election Study (CES) surveys spanning 2016-2024. The original CES data included open-ended categories that respondents could elaborate on in text fields. This recoding effort systematically categorizes these responses into structured variables that can be merged with the original CES data.

2 Dataset Description

2.1 Files Included

The dataset includes 12 CSV files total:

Combined Files (2016-2024):

- **Religion Combined**: All years of recoded open-ended religion responses
- **Race Combined**: All years of recoded open-ended race/ethnicity responses

Individual Year Files:

- **Religion by Year**: Separate files for each survey year (2016-2024)
- **Race by Year**: Separate files for each survey year (2016-2024)

All datasets can be merged with original CES data using `caseid` and `year` (if necessary) as matching keys. Users can choose between the combined files for multi-year analyses or individual year files for year-specific studies.

3 Religion Data Structure

3.1 Variables

Variable	Description
`year`	Survey year (2016-2024)
`caseid`	Unique respondent identifier within year
`religpew_new`	Primary religion category based on first mentioned religion

`religpew_t`	Original open-ended text response
`religpew_evangelical`	Binary indicator (1/NA) if response mentioned evangelical content
`religpew_spiritual`	Binary indicator (1/NA) if response mentioned spiritual content
`religpew_refuse`	Binary indicator (1/NA) if respondent refused to answer
`religpew_1` through `religpew_12`	Binary indicators (1/NA) for all religions

3.2 Religion Category Codes

The recoding follows the standard CES religion classification with numeric codes:

- 1 = Protestant
- 2 = Roman Catholic
- 3 = Mormon
- 4 = Eastern or Greek Orthodox
- 5 = Jewish
- 6 = Muslim
- 7 = Buddhist
- 8 = Hindu
- 9 = Atheist
- 10 = Agnostic
- 11 = Nothing in particular
- 12 = Something else

3.3 Multiple Religion Handling

When respondents mentioned multiple religious affiliations:

- `religpew_new` captures the first/primary religion mentioned
- `religpew_1` through `religpew_12` receive binary coding (1) for each additional religion category mentioned
- Unmention categories receive NA values

4 Race Data Structure

4.1 Variables

Variable	Description
`caseid`	Unique respondent identifier within year
`race`	Numeric code corresponding to CES race categories (7 = Other)
`race_other`	Original open-ended text response
`year`	Survey year (2016-2024)
`race_t`	Recoded race category (NA values remain coded as "Other")
`hispanic_t`	Binary indicator (1/NA) if response mentioned Hispanic/Latino content in addition to another identity
`refuse_t`	Binary indicator (1/NA) if respondent refused to answer

4.2 Race Category Codes

The recoding follows the standard CES race/ethnicity classification with numeric codes:

- 1 = White
- 2 = Black
- 3 = Hispanic
- 4 = Asian
- 5 = Native American
- 6 = Two or more races
- 7 = Other
- 8 = Middle Eastern

5 Data Integration

5.1 Merging with Original CES Data

To incorporate these recoded variables with the original CES datasets:

1. Merge both CSV files with CES data using `caseid` and `year` (if necessary)
2. For religion: Create a composite variable by coalescing the original `religpew` with `religpew_new`
3. For race: Create a composite variable by coalescing the original `race` with `race_t`
4. Retain binary indicators (`religpew_evangelical`, `religpew_spiritual`, `hispanic_t`, etc.) as separate variables for nuanced identity analysis

5.2 Working with Binary Indicators

The binary indicator variables (coded as 1 or NA) provide additional layers of identity information:

- **Religion**: Use `religpew_evangelical` and `religpew_spiritual` to identify respondents who mentioned these characteristics regardless of their primary religious category
- **Race**: Use `hispanic_t` to identify respondents who mentioned Hispanic/Latino identity in addition to other racial categories
- **Multiple affiliations**: Use `religpew_1` through `religpew_12` to identify respondents with multiple religious affiliations beyond their primary category

6 Contact Information

For questions about the recoding methodology or data structure, contact caroline.soler@tufts.edu.