Guanzheng Sun
Spring 2024
Codebook and Data Set Assignment

- *What is the title/name of your data set?*
  National Asian American Survey (2008, 2016 Pre-Election, and 2016 Post-Election).

- *What is the format of your data file?: SPSS, Stata, Minitab, etc.*
  All 3 data files are in SPSS format.

- *Who (or what) were the investigators who compiled the data?*
  Subramanian Karthick, Jane Junn, Taeku Lee, and Janelle Wong

- *Where (or from what site) you did you access the data?*
  ICPSR.

- *Details of data collection (e.g., sample or universe, nature of sample, cases, and other relevant methodological details [e.g., when collected, if survey how administered]).*

  2008:
  Samples, 5,159 completed telephone interviews of self-identified AAPI residents of the United States. Interviewing began on August 12, 2008, and ended on October 29, 2008. Covered topics such as political behavior and attitudes, personal experiences in immigration to the United States, attitudes toward government, politics, political issues, extent of political involvement, party affiliation, sources of political information, voting behavior, health and financial status, racial and ethnic identification, and religious and ethnic social networks.

  2016 Pre:
  Samples, 4787 cases of nationally representative data from telephone interviews of adult U.S. residents self-identified as AAPI. Covered questions about immigrant background, social identities, social attitudes, political behavior, and policy attitudes.

  2016 Post:
  Samples, 6448 cases of nationally representative data from telephone interviews of adult U.S. residents self-identified as AAPI. Covered questions about immigrant background, social identities, social attitudes, political behavior, and policy attitudes.

- *Variables analyzed, how these variables were operationalized. (Note: If you are drawing from dataset with a large number of variables not related to your research question in any reasonable way, then include only those select variables that are relevant in the codebook – and in the data set too.)*

  My major research inquiries for this research project include:

1. Are Chinese Americans becoming "redder" in years? How about compared to Asian Americans?
2. If not, what do their party affiliations look like? How about compared to Asian Americans?
3. What are the major issues identified by Chinese Americans?
4. How are these issues associated/affecting their Presidential voting preference and party affiliation?

I also have two more specific questions for Chinese Americans:

a. How much does education/school quality matter to Chinese Americans? Since they are the ones covered most in the news opposing affirmative action.
b. In 2016, how much were Asian Americans/Chinese Americans choosing the Republican Party? Or they were more choosing Trump?

Something came up in the data analyses process also:

a. Why are Chinese Americans more reserved in expressing their political judgements/opinions/expressions? Are they interested in politics at all?

Thus, the variables operationalized based on these inquiries included:

General Party Affiliation & Presidential Election

	2008	2016-Pre	2016-Post
Ethnicity	QA1	RETHNIC	RETHNICX
Party General	QD1, QD1C	PID4	Q10.0A, Q10.0D
Presidential	QC3	Q4_5A	Q2_4A

Factors Affecting Party Affiliation:

	2008	2016-Pre	2016-Post
Place of Birth	QA4	S9	Q1_3A
Interest in Politics	QB1	Q3_1	Q2_1A
Pivotal Issues	QB5A	Q6_1D	N/A

Issues:

	2008	2016-Pre	2016-Post
Demonstration/Rally	QC15_7	Q5_1_05	N/A
Home Country Political Activity	QC20	Q6_4A	N/A

LGBTQ+	QG6	N/A	Transgender
Police	N/A	Q6_4A	N/A
Education	N/A	N/A	Q5_7B

- *If using recoded or computing new variables, detail on strategy.*

For 2008, 2016 Pre, and 2016 Post-Election dataset, I recoded several variables.

For 2016 Pre, the variables I recoded are:

Pivotal_Issues: Recoded from Q6.1D Which issue is the most important to you PERSONALLY?), I filtered issues pivotal to Asian Americans in 2016 (1.00 = Economy/Jobs, 2.00 = Education, 4.00 = Foreign Policy, 8.00 = Immigrants, 9.00 = Racism, 10.00 = LGBTQ+, 12.00 = Taxes, the numbers here are values given by original researchers.

Major_Asians: I selected 6 major Asian American Groups to represent the entire Asian American group (since the original RETHNIC variable recording ethnicity compositions of respondents includes 60 ethnicities, I decided to only keep 6 most representative groups according to census bureau's statistics: Chinese, Filipino, Indian, Japanese, Korean, and Vietnamese) They are used to compare and contrast with the issue preference/party affiliation/presidential votes with Chinese American.

Chinese: Filtering out Chinese survey respondents only

Party_Affiliation: A recoded variable from PID_4 -- Partisan Identification 3 Categories (D, R, I), I only kept 1.00 – Democratic and 2.00 – Republican.

Presidential_Preference: A recoded variable showcasing survey respondent's preference for one specific presidential candidate in the upcoming 2016 presidential election. I filtered out the various small party candidates/no response/missing responses to only compare the contest between Clinton and Trump.

For 2016 Post-Election, the variables I recoded are:

Same manner of recoding for Presidential_Preference (although in this case it becomes Presidential_Vote since the respondents have already voted), Party_Affiliation, Major_Asians, and Chinese.

Unique new recoded variables are:

School_Quality: reflects the respondent's opinion on the issue of whether school quality matters to them.

Equal_Rights: reflects the respondent's opinion on the issue of equal rights, is it right or wrong?

Transgender: reflects the respondent's opinion on transgender issues.

For 2008, recoded variables remain same in name but coded differently according to difference in original coding:

For Presidential_Preference, 1.00 = McCain and 2.00 = Obama (Plan to adjust this in upcoming 2 weeks). Similarly for Party_Affiliation, 1.00 = Republican and 2.00 = Democratic.

For Pivotal_Issues, 1 = Iraq War, 2-3 = Economy/Jobs, 6 = Education, 7-8 = Immigrants, 12-13 = Racism.

- *If you weighted the data, variable used for weighting – and generally on what basis data were data weighted (e.g., gender, age)?*

  After weighing the data using the weight variable recommended by the original data researchers, we found out that it produces a gigantic amount of cases for each variable, making it unfair and unrealistic to produce any meaningful analyses since any of the results would be statistically significant.

- *Include a link that takes me to the basic data you started with. You don't have to submit your dataset at this point, though I might down the road ask for your dataset.*
  Included in the citations below.

Citations:

Bureau, US Census. "20.6 Million People in the U.S. Identify as Asian, Native Hawaiian or Pacific Islander." *Census.gov*. https://www.census.gov/library/stories/2022/05/aanhpi-population-diverse-geographically-dispersed.html (May 3, 2024).

Ramakrishnan, S. Karthick (Subramanian Karthick), Jane Junn, Taeku Lee, and Janelle Wong. 2011. "National Asian American Survey (NAAS), [United States], 2008: Version 2." doi:10.3886/ICPSR31481.V2.

Ramakrishnan, S. Karthick (Subramanian Karthick), Jennifer Lee, Taeku Lee, and Janelle Wong. 2018. "National Asian American Survey (NAAS) Pre-Election Survey, [United States], 2016: Version 1." doi:10.3886/ICPSR37024.V1.

Ramakrishnan, S. Karthick (Subramanian Karthick), Jennifer Lee, Taeku Lee, and Janelle Wong. 2020. "National Asian American Survey (NAAS) Post-Election Survey, [United States], 2016: Version 1." doi:10.3886/ICPSR37380.V1.