Codebook - Latinobarometro-2018

Christopher B.

2024-02-26

Contents

1. **Overview of Data**

2. **Sources and Methodology**

3. **Note on Missing Values**

4. **Presentation of Variables**

Overview of Data

Description:

Latinobarómetro is a public opionon study that conducts over 20,000 interviews across 18 Latin American countries each year, representing over 600 million inhabitants. The Latinobarómetro Corporation (Corporación Latinobarómetro) is a non-profit NGO based in Santiago, Chile and is the sole producer and publisher of the Latinobarómetro survey data. This data covers public opinions, attitudes, values, and behaviors, including data on democracy, media, economic and political confidence, and on. Their data spans from 1995 to present.

This dataset, *latinobarometro_2018.RData*, limits the sample for the following countries:

- Colombia

- Ecuador

- El Salvador

- Guatemala

- Honduras

- Nicaragua

- Paraguay

- Peru

and delivers the following cleaned variables:

- c_name: country name

- source: "Latinobarometro"

- r_id: respondent id

- r_year: year of interview

- r_month: month of interview

- r_swt: survey weight

- r_female: binary indicator for female

- p_ideology: left-right ideological self-placement

- p_interest: interest in politics

- p_economy: evaluation of national economy, past year

- 4 measures of satisfaction with and/or preference for democracy

Several variables were created from existing variables within the *Latinobarometro_2018_Eng_Stata_v20190303.dta* dataset.

The dataset was created in part for a course at American University's School of International Service (SIS-750-007 Data Analysis) for potential use in research and publication.

Resources (Hyperlinks):

- UNESCO Website Description (English)

- Latinobarómetro Website (Spanish)

- Latinobarómetro Website (English, Google Translated)

- UNDP Website Results of Latinobarómetro 2020 (Spanish)

- Rutger's University Library Website Description (English)

Sources and Methodology

Description:

To create the Latinobarometro-2018 dataset, I first downloaded the 2018 English data *Latinobarometro_2018_Eng_Stata_v20190303.dta* from the following link: https://www.latinobarometro.org/latContents.jsp?Idioma=0&CMSID=Datos&Idioma=0&CMSID=Datos. I also downloaded their English 2018 codebook from the following link: https://www.latinobarometro.org/latContents.jsp.

To limit the sample to just Colombia, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Paraguay, and Peru, I filtered the dataset by each country's IDENPA number.

To deliver the requested variables, I either transposed, wrangled, or created variables from the *Latinobarometro_2018_Eng_Stata_v20190303.dta* dataset. Specifically, I transposed the following variables from the original dataset:

- Country (IDENPA)

- Respondent ID (NUMENTRE)

- Year of Interview (NUMINVES)

- Month of Interview (MESREAL)

- Survey Weight (WT)

- Female (SEXO)

- Left-Right Ideological Self-Placement (P22ST)

- Evaluation of the National Economy (P6STGBSC)

- Support for Democracy (P12STGBS)

- Satisfaction with Democracy (P13STGBS.A)

- Level of Democracy (P18GBS)

- Belief in Democracy as the Best System (P24ST)

I wrangled or created many of these variables to get variable and value labels that were more conducive to the dataset. Here are the following wrangling steps after selecting and limiting the data:

- created a variable c_name for Country. Changed the c_name values (from IDENPA) to a character, then renamed all observations in c_name from the IDENPA number to the country's name (ex. 170 –> Colombia)

- created the Source variable by mutating a variable (source) that holds the string "Latinobarometro".

- created the r_id variable and cleaned it by subtracting the unique value added to it (taken originally from its IDENPA score) so that the r_id would correspond better with each country (instead of being a unique value). For example, 1700000 had been added to the r_id of in the original dataset to render it a unique value. I simply subtracted 1700000 so that r_id was not unique to each country.

- created r_year and set it to numeric

- created r_month and set it to numeric

- created r_swt and set it to numeric

- reconfigured Female (r_female) to be 0 if Non-Female and 1 if Female, whereas it had been 1 if Male and 2 if Female.

- added variable labels using the *apply_labels()* function and *mutate(labelled())* functions

I created the Political Interest variable using the P19ST.A, P19ST.B, P19ST.C, P19ST.D, P19ST.E, P19ST.F, P19ST.G, P19ST.H, P19NC.I, P19NC.J, P19F.K, and P19ST.L variables from the *Latinobarometro_2018_Eng_Stata_v20190303.dta* dataset. It is an incredibly rough proxy for political interest, but I was unable to find any other sufficient proxies within the *Latinobarometro_2018_Eng_Stata_v20190303.dta* dataset.

I ordered and arranged the dataset in the following order: Country, Source, Respondent ID, Year of Interview, Month of Interview, Survey Weight, Female, Left-Right Ideological Self-Placement, Political Interest, Evaluation of the National Economy, Support for Democracy, Satisfaction with Democracy, Level of Democracy, Belief in Democracy.

Note on Missing Values

The *Latinobarometro_2018_Eng_Stata_v20190303.dta* dataset was coded with several values. For instance, in many cases, the following codes were used:

- -8: NA

- -5: Missing

- -4: Not Asked

- -3: Not Applicable

- -2: No Answer/Refused

- -1: Don't Know

My dataset recoded out the -8 and -5 for NA values and kept the original code of -4, -3, -2, -1.

Presentation of Variables

Country

Variable Name: c_name

Variable Type: character

Description: Country in which the interview took place. (Source: IDENPA in Latinobarometro 2018.)

c_name	freq	perc
Colombia	1200	13.6
Ecuador	1200	13.6
El Salvador	1000	11.4
Guatemala	1000	11.4
Honduras	1000	11.4
Nicaragua	1000	11.4
Paraguay	1200	13.6
Peru	1200	13.6

Source

Variable Name: source

Variable Type: character

Description: interview data source (from: Latinobarómetro 2018). (Source: Created.)

source	freq	perc
Latinobarometro	8800	100

Respondent ID

Variable Name: r_id

Variable Type: numeric

Description: Respondent ID (Per Country). (Source: NUMENTRE in Latinobarometro 2018.)

min	max
1	1200

Year of Interview

Variable Name: r_year

Variable Type: numeric

Description: Year of Interview (2018). (Source: NUMINVES in Latinobarometro 2018.)

Month of Interview

Variable Name: r_month

Variable Type: numeric

Description: Month of Interview. (Source: MESREAL in Latinobarometro 2018.)

Month	Label	freq	perc
6	June	4451	50.6
7	July	4346	49.4
8	August	3	0.0

Survey Weight

Variable Name: r_swt

Variable Type: numeric

Description: Survey Weight. (Source: WT in Latinobarometro 2018.)

min	mean	median	max
0.452	1	1	4.454

Female

Variable Name: r_female

Variable Type: numeric

Description: Binary Indicator for Female. (Source: SEXO in Latinobarometro 2018.)

min	mean	median	max
0	0.519	1	1

Left-Right Ideological Self-Placement

Variable Name: p_ideology

Variable Type: numeric

Description: Self-Positioning in Left-Right Scale (1-10). (Source: P22ST in Latinobarometro 2018.)

min	mean	median	max
0	5.316	5	10

NA Values:

- -1 = Don't Know
- -2 = No Answer/Refused

- -3 = Not Applicable

- -4 = Not Asked

- *NA* = NA

Political Interest

Variable Name: p_interest

Variable Type: numeric

Description: Level of interest in politics is based on the number of political sources of information (from original Latinobarometro 2018 dataset — P19ST.A through P19ST.M variables) through which the person learns about politics. There are a total of 12 variables from the Latinobarometro data: P19ST.A (family), P19ST.B (friends), P19ST.C (co-workers), P19ST.D (classmates), P19ST.E (radio), P19ST.F (newspapers/magazines), P19ST.G (electronic media / internet), P19ST.H (television), P19NC.I (facebook), P19NC.J (twitter), P19F.K (youtube), and P19ST.L (other). The variable P19ST.M (none) was not used to create p_interest. Unlike the 2020 and 2023 datasets, there was not a proxy indicator for political interest within the Latinobarometro 2018 dataset. This variable therefore acts as a very rough proxy indicator for interest in politics. The idea is that people who are more interested in politics may be more likely to be informed via more avenues of political information. (Source: Created using P19ST.A, P19ST.B, P19ST.C, P19ST.D, P19ST.E, P19ST.F, P19ST.G, P19ST.H, P19NC.I, P19NC.J, P19F.K, P19ST.L, P19ST.M in Latinobarometro 2018.)

min	mean	median	max
0	3.711	3	12

NA Values:

- -1 = Don't Know

- -2 = No Answer/Refused

- -3 = Not Applicable

- -4 = Not Asked

- *NA* = NA

Evaluation of the National Economy

Variable Name: p_economy

Variable Type: numeric

Description: Current Economic Situation of the Country. (Source: P6STGBSC in Latinobarometro 2018.)

min	mean	median	max
1	3.392	3	5

NA Values:

- -1 = Don't Know
- -2 = No Answer/Refused
- -3 = Not Applicable
- -4 = Not Asked
- *NA* = NA

Support for Democracy

Variable Name: p_democracy_preferable

Variable Type: numeric

Description: Support for democracy. 1. Democracy is preferable to any other kind of government, 2. Under some circumstances, an authoritarian government can be preferable to a democratic one, 3. For people like me, it does not matter whether we have a democratic or a non-democratic regime. (Source: P12STGBS in Latinobarometro 2018.)

min	mean	median	max
1	1.882	2	3

NA Values:

- -1 = Don't Know
- -2 = No Answer/Refused
- -3 = Not Applicable
- -4 = Not Asked
- *NA* = NA

Satisfaction with Democracy

Variable Name: p_democracy_satisfaction

Variable Type: numeric

Description: Satisfaction with Democracy. 1. Very Satisfied, 2. Rather Satisfied, 3. Not Very Satisfied, 4. Not At All Satisfied. (Source: P13STGBS in Latinobarometro 2018.)

min	mean	median	max
1	2.984	3	4

NA Values:

- -1 = Don't Know
- -2 = No Answer/Refused
- -3 = Not Applicable
- -4 = Not Asked
- *NA* = NA

Level of Democracy in (Country)

Variable Name: p_democracy_level

Variable Type: numeric

Description: How would you characterize Democracy in your country? (¿Cómo diría Ud. que es la democracia en su país?). 1. A Full Democracy, 2. A Democracy With Little Problems, 3. A Democracy With Big Problems, 4. Not A Democracy (Source: P18GBS in Latinobarometro 2018.)

min	mean	median	max
1	2.852	3	4

NA Values: - 5 = I don't understand what a democracy is

- -1 = Don't Know
- -2 = No Answer/Refused
- -3 = Not Applicable
- -4 = Not Asked
- NA = NA

Belief in Democracy as the Best System

Variable Name: p_democracy_best

Variable Type: numeric

Description: Agreement/Disagreement: Democracy may have problems but it is the best system of government. 1. Strongly Agree, 2. Agree, 3. Disagree, 4. Strongly Disagree (Source: P24ST in Latinobarometro 2018.)

min	mean	median	max
1	2.241	2	4

NA Values:

- -1 = Don't Know
- -2 = No Answer/Refused
- -3 = Not Applicable
- -4 = Not Asked
- NA = NA