**Codebook for the Regional Study of Venezuelan Migration**

**Colombia**

Data on Venezuelan migrants in Colombia comes from the Migration Pulse Survey, conducted across four rounds between 2021 and 2022. This survey is conducted in four rounds between 2021 and 2022 and uses computer-assisted telephone interviews (CATI) to collect longitudinal data from Venezuelan migrants aged 15 years and above. The survey aims to represent the Venezuelan migrant population residing in Colombia and includes a sample size of approximately 4,200 households, with at least one migrant member in each household. The identification of these households is carried out using the migration module of the Great Integrated Household Survey (GEIH, Spanish acronym).

Implemented by the National Administrative Department of Statistics (DANE) in collaboration with the World Bank and Universidad del Rosario, the Migration Pulse Survey delves into various aspects of Venezuelans' lives, including living conditions, migration status, socio-demographics, and employment. Venezuelan migrants are depicted using the fourth wave of the survey[[1]](#footnote-1), collected between March and April of 2022, with a sample of 6,176 individuals resulting in an expanded coverage of 1,554,698 individuals.

The characterization of Colombian host communities relies on data from the 2021 Great Integrated Household Survey (GEIH, Spanish acronym). As the primary household survey employed to monitor labor market outcomes in the country, the GEIH is conducted by the National Administrative Department of Statistics (DANE). It serves as a comprehensive source of information, capturing sociodemographic and labor market characteristics on a national scale.

The GEIH incorporates various modules designed to address specific issues, including the migration module, which is the foundation for the Migration Pulse Survey. In 2021, the GEIH featured a robust annual sample of around 232,000 households, providing representation across the 24 capital cities and metropolitan areas (including Bogotá), rural and urban geographical zones, and all 23 departments. The sample of hosting communities is exclusively drawn from urban households. The unweighted host community sample comprises 447,888 Colombians, offering insights that extend to a broader population of 26,370,254 individuals.

|  |  |
| --- | --- |
| **Label** | **Variable List** |
| **Identifiers:** |  |
| Ponderation weight | weight |
| Year | year |
| Survey | survey |
| Sample | samp |
| Wave  | wave |
| Region code | depcode |
| **Demographics:** |  |
| Sex | sex |
| Female | female |
| Male | male |
| Age | age |
| Household size | hh\_size |
| Marital status | marital\_status |
| Married | married |
| Years of education | edu\_years |
| Education level | edu\_level |
| **Employment:** |  |
| Employed | employed |
| Unemployed | unemployed |
| Inactive | inactive |
| Monthly income (US$ PPP17 - censored 2000) | cen\_minc |
| **Migration:** |  |
| Do you have a residence permit? | entry\_register |
| Have you homologated your title? | homologate |
| Why haven't you homologated your title? | reason\_nohomologate |
| Transport mean mostly used in migration journey | transport |
| With whom have you traveled from Venezuela? | company |
| Have you sent remittances abroad? | remittances |
| Residency plans | plan\_residency |
| Have you experienced discrimination or unjust treatment? | discriminated |
| (Integration in) Religious groups | integration\_2\_1 |
| (Integration in) Sports clubs or groups | integration\_2\_2 |
| (Integration in) Community-specific groups (youth, ethnicity, LGBT, etc.) | integration\_2\_3 |
| (Integration in) Other | integration\_2\_4 |
| (Integration in) None | integration\_2\_5 |

**Survey Identifiers**

|  |  |
| --- | --- |
| **Variable Name:** | **weight** |
| **Variable Label:** | Ponderation weight |
| **Description:** | The variable indicates the ponderation weight of each observation to be used as a frequency weight |
| **Source:** | *Migrants:* Migration Pulse, wave 1: FACTOR, wave 2: FEX, wave 4: FEX. *Hosts:* GEIH, variable fex\_c18weight = fex\_c18/12 |

|  |  |
| --- | --- |
| **Variable Name:** | **year** |
| **Variable Label:** | Year |
| **Description:** | The variable indicates the year in which the survey was conducted |

|  |  |
| --- | --- |
| **Variable Name:** | **survey** |
| **Variable Label:** | Survey |
| **Description:** | The variable indicates the survey from where the data comes from |
| **Format** | 1 = Migration Pulse2 = GEIH |

|  |  |
| --- | --- |
| **Variable Name:** | **samp** |
| **Variable Label:** | Sample |
| **Description:** | The variable indicates whether the observation refers to a Venezuelan migrant or a Colombian host. |
| **Format** | 1 = Migrants2 = Hosts |

|  |  |
| --- | --- |
| **Variable Name:** | **wave** |
| **Variable Label:** | Wave |
| **Description:** | The variable indicates the wave of the survey from where the data comes from |
| **Format** | wave = 1, 2, or 4 |
| **Variable Name:** | **depcode** |
| **Variable Label:** | Region code |
| **Description:** | The variable indicates the region of the respondent |
| **Format** | 5 = ANTIOQUIA 8 = ATLANTICO11 = BOGOTA13 = BOLIVAR15 = BOYACA17 = CALDAS18 = CAQUETA19 = CAUCA20 = CESAR23 = CORDOBA25 = CUNDINAMARCA27 = CHOCO41 = HUILA44 = LA GUAJIRA47 = MAGDALENA50 = META52 = NARIÑO54 = N. DE SANTANDER63 = QUINDIO66 = RISARALDA68 = SANTANDER70 = SUCRE73 = TOLIMA76 = VALLE DEL CAUCA |

**Demographics**

|  |  |
| --- | --- |
| **Variable Name:** | **sex** |
| **Variable Label:** | Sex |
| **Description:** | The variable indicates whether the observation refers to a male or a female respondent |
| **Format:** | 1 = Male2 = Female |
| **Source:** | *Migrants:* Migration Pulse, wave 4, variable Ide9*Hosts:* GEIH, variable p6020 |

|  |  |
| --- | --- |
| **Variable Name:** | **female** |
| **Variable Label:** | Female |
| **Description:** | The variable indicates whether the observation refers to a female respondent |
| **Format:** | 0 = Male1 = Female |
| **Source:** | *Migrants:* Migration Pulse, wave 4, variable Ide9*Hosts:* GEIH, variable p6020 |

|  |  |
| --- | --- |
| **Variable Name:** | **male** |
| **Variable Label:** | Male |
| **Description:** | The variable indicates whether the observation refers to a male respondent |
| **Format** | 0 = Female1 = Male |
| **Source:** | *Migrants:* Migration Pulse, wave 4, variable Ide9*Hosts:* GEIH, variable p6020 |

|  |  |
| --- | --- |
| **Variable Name:** | **age** |
| **Variable Label:** | Age |
| **Description:** | The variable indicates the age of the respondent in years. Limited to individuals between 18 and 100 years of age |
| **Format** | Years |
| **Source:** | *Migrants:* Migration Pulse, wave 4, variable Ide10*Hosts:* GEIH, variable p6040 |

|  |  |
| --- | --- |
| **Variable Name:** | **hh\_size** |
| **Variable Label:** | Household size |
| **Description:** | The variable indicates the number of household members |
| **Source:** | *Migrants:* Migration Pulse, wave 1, variable Ide12*Hosts:* GEIH, variable p6008 (use only the response of the head of household) |

|  |  |
| --- | --- |
| **Variable Name:** | **marital\_status** |
| **Variable Label:** | Marital status |
| **Description:** | The variable indicates the marital status of the respondent |
| **Format** | 1 = Single2 = Married/Cohabitation3 = Other |
| **Source/Algorithm:** | *Hosts:* GEIH, variable p6070marital\_status = 1 if p6070 = 6marital\_status = 2 if p6070 = 1, 2, or 3marital\_status = 3 if p6070 = 4, or 5 |

|  |  |
| --- | --- |
| **Variable Name:** | **married** |
| **Variable Label:** | Married |
| **Description:** | The variable indicates whether the respondent is married |
| **Format** | 0 = No1 = Yes |
| **Source/Algorithm:** | *Migrants:* Migration Pulse, wave 1, variable mi7married = 1 if mi7 ≠ 5 |

|  |  |
| --- | --- |
| **Variable Name:** | **edu\_years** |
| **Variable Label:** | Years of education |
| **Description:** | The variable indicates the number of years of education completed at the time of the survey |
| **Format** | Years |
| **Source/Algorithm:** | *Migrants:* Migration Pulse, wave 1, variable ml1edu\_years = 0 if ml1 = 1edu\_years = 7 if ml1 = 2edu\_years = 13 if ml1 = 3edu\_years = 16 if ml1 = 4edu\_years = 18 if ml1 = 5edu\_years = 23 if ml1 = 6*Hosts:* GEIH, variable esc |

|  |  |
| --- | --- |
| **Variable Name:** | **edu\_level** |
| **Variable Label:** | Education level |
| **Description:** | The variable indicates the highest level of education completed by the respondent |
| **Format** | 0 = None1 = Primary2 = Secondary3 = Tertiary |
| **Source/Algorithm:** | *Migrants:* Migration Pulse, wave 1, variable ml1edu\_level = 0 if ml1 = 1edu\_level = 1 if ml1 = 2edu\_level = 2 if ml1 = 3edu\_level = 3 if ml1 = 4, 5, or 6*Hosts:* GEIH, variable p6210edu\_level = 0 if p6210 = 1edu\_level = 1 if p6210 = 2, or 3edu\_level = 2 if p6210 = 4, or 5edu\_level = 3 if p6210 = 6 |

**Employment**

|  |  |
| --- | --- |
| **Variable Name:** | **employed** |
| **Variable Label:** | Employed |
| **Description:** | The variable indicates whether the respondent has been employed in the last week for individuals between 18 and 65 years of age |
| **Format** | 0 = No1 = Yes |
| **Source/Algorithm:** | *Migrants:* Migration Pulse, wave 4, variable ml5employed = 1 if ml5 = 1*Hosts:* GEIH, variables p6240, p6250, p6260employed = 1 if p6240, p6250, or p6260 = 1 |

|  |  |
| --- | --- |
| **Variable Name:** | **unemployed** |
| **Variable Label:** | Unemployed |
| **Description:** | The variable indicates whether the respondent was unemployed but looking in the last 4 weeks for individuals between 18 and 65 years of age |
| **Format** | 0 = No1 = Yes |
| **Source/Algorithm:** | *Migrants:* Migration Pulse, wave 4, variable ml5employed = 1 if ml5 = 2*Hosts:* GEIH, variable p6240employed = 1 if p6240 = 2 |

|  |  |
| --- | --- |
| **Variable Name:** | **inactive** |
| **Variable Label:** | Inactive |
| **Description:** | The variable indicates whether the respondent was not working last week nor looking for a job in the last 4 weeks for individuals between 18 and 65 years of age |
| **Format** | 0 = No1 = Yes |
| **Source/Algorithm:** | *Migrants:* Migration Pulse, wave 4inactive = 1 if unemployed = 0 and employed = 0*Hosts:* GEIHinactive = 1 if unemployed = 0 and employed = 0 |

|  |  |
| --- | --- |
| **Variable Name:** | **cen\_minc** |
| **Variable Label:** | Monthly income (US$ PPP17 - censored 2000) |
| **Description:** | The variable represents the monthly earned income of individuals between 18 and 65 years of age  |
| **Format** | The value of this variable ranges from 26.4 to 2,000.This variable is in dollars at the PPP2017. |
| **Source:** | *Migrants:* Migration Pulse, wave 1, variables ml324 ml325a ml325b (merged from wave 3)cen\_minc = (ml324 if ml324>100) + (ml325a if ml325b == 2)Divide by PPP factor 2017 = 1467.28, and censor at 2000 |

**Migration**

|  |  |
| --- | --- |
| **Variable Name:** | **entry\_register** |
| **Variable Label:** | Do you have a residence permit? |
| **Description:** | The variable indicates whether the respondent has a residence permit |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 4, variable mi4\_9entry\_register = 1 if mi4\_9 = 1 |

|  |  |
| --- | --- |
| **Variable Name:** | **homologate** |
| **Variable Label:** | Have you homologated your title? |
| **Description:** | The variable indicates whether the respondent has homologated their title in Colombia |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 1, variable ml2  |

|  |  |
| --- | --- |
| **Variable Name:** | **reason\_nohomologate** |
| **Variable Label:** | Why haven't you homologated your title? |
| **Description:** | The variable indicates the reasons why the respondent has not homologated it titles in Chile |
| **Format** | 1 = Don’t know the process2 = Lack of necessary documents3 = Money4 = No interest or need5 = Other |
| **Source:** | Migration Pulse, wave 1, variable ml3reason\_nohomologate = 1 if ml3 = 1reason\_nohomologate = 2 if ml3 = 2reason\_nohomologate = 3 if ml3 = 3reason\_nohomologate = 4 if ml3 = 4, or 5reason\_nohomologate = 5 if ml3 = 6 |

|  |  |
| --- | --- |
| **Variable Name:** | **transport** |
| **Variable Label:** | Transport means mostly used in migration journey |
| **Description:** | The variable indicates what the mostly used transportation means  |
| **Format** | 1 = By foot2 = Private transport3 = Public transport |
| **Source:** | Migration Pulse, wave 1, variable mi6 |

|  |  |
| --- | --- |
| **Variable Name:** | **company** |
| **Variable Label:** | With whom have you traveled from Venezuela? |
| **Description:** | The variable indicates the people the respondent was accompanied by during their migration journey |
| **Format** | 1 = Relatives2 = Not relatives3 = Mixed4 = Alone |
| **Source:** | Migration Pulse, wave 1, variable mi5company = 1 if mi5 = 1company = 2 if mi5 = 2company = 4 if mi5 = 3 |

|  |  |
| --- | --- |
| **Variable Name:** | **remittances** |
| **Variable Label:** | Have you sent remittances abroad? |
| **Description:** | The variable indicates whether the respondent has sent remittances abroad |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 1, variable ir7 |

|  |  |
| --- | --- |
| **Variable Name:** | **plan\_residency** |
| **Variable Label:** | Residency plans |
| **Description:** | The variable indicates whether the respondent plans to stay in Peru |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 1, variable mi23 |

|  |  |
| --- | --- |
| **Variable Name:** | **discriminated** |
| **Variable Label:** | Have you experienced discrimination or unjust treatment? |
| **Description:** | The variable indicates whether the respondent has felt discriminated in Peru |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 2, variable vi4 |

|  |  |
| --- | --- |
| **Variable Name:** | **integration\_2\_1** |
| **Variable Label:** | (Integration in) Religious groups |
| **Description:** | The variable indicates whether the respondent has participated in religious groups as part of integration efforts in Colombia (participation in this activity is not mutually exclusive with other activities) |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 2, variable ei3\_\_2 |

|  |  |
| --- | --- |
| **Variable Name:** | **integration\_2\_2** |
| **Variable Label:** | (Integration in) Sports clubs or groups |
| **Description:** | The variable indicates whether the respondent has participated in Sports clubs or groups as part of integration efforts in Colombia (participation in this activity is not mutually exclusive with other activities) |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 2, variable ei3\_\_1 |

|  |  |
| --- | --- |
| **Variable Name:** | **integration\_2\_3** |
| **Variable Label:** | (Integration in) Community-specific groups (youth, ethnicity, LGBT, etc.) |
| **Description:** | The variable indicates whether the respondent has participated in community-specific groups (youth, ethnicity, LGBT, etc.) as part of integration efforts in Peru (participation in this activity is not mutually exclusive with other activities) |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 2, variables ei3\_\_3, ei3\_\_4, ei3\_\_5, ei3\_\_6, ei3\_\_7, ei3\_\_8, ei3\_\_9integration\_2\_3 = 1 if ei3\_\_3, ei3\_\_4, ei3\_\_5, ei3\_\_6, ei3\_\_7, ei3\_\_8, or ei3\_\_9 = 1 |

|  |  |
| --- | --- |
| **Variable Name:** | **integration\_2\_4** |
| **Variable Label:** | (Integration in) Other |
| **Description:** | The variable indicates whether the respondent has participated in other activities as part of integration efforts in Colombia (participation in this activity is not mutually exclusive with other activities) |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 2, variable ei3\_\_10 |

|  |  |
| --- | --- |
| **Variable Name:** | **integration\_2\_5** |
| **Variable Label:** | (Integration in) None |
| **Description:** | The variable indicates whether the respondent has not participated in any activity as part of integration efforts in Colombia |
| **Format** | 0 = No1 = Yes |
| **Source:** | Migration Pulse, wave 2, variable ei3\_\_11 |

1. Information from the first (July-August 2021) and second (October-November 2021) waves was also used depending on availability. [↑](#footnote-ref-1)