

Metadata

| | |
|-------------|-----------------------------|
| Version | 1 |
| Created by | Translators without Borders |
| Uploaded on | Apr-21 |

| | |
|----------------------------|--|
| Notes and caveats | <p>All data is drawn from government survey results and is subject to any associated limitations or distortions present in the source data.</p> <p>Literacy was measured as the ability to read and write in any language.</p> <p>As a data protection measure, the language dataset provider, IPUMS, has merged various second administrative units in order to provide spatio-temporally harmonized units that meet their 20,000 population threshold required for public dissemination of the data. This has resulted in combined figures for different zones in the dataset. Tounifili and Mankountan, for example, have been combined into one row in this dataset and the language data can only apply across these combined zones. Attempts to disaggregate these locations would likely result in decreased data accuracy.</p> <p>Languages with population shares less than 0.01% or under 1,000 people (whichever is greater) have been aggregated into the “Other” field.</p> <p>All decimal values have been rounded to a maximum of 3 decimal places. As a result language shares may not total 100%.</p> |
| Copyright and terms of use | <p>Empty values represent non-existent data and should not be treated as zero values.</p> <p>Data is available under an Attribution NonCommercial ShareAlike 4.0 International license (CC BY NC SA 4.0)</p> <p>You are free to share and adapt the data subject to requirements for attribution and non commercial use. Any derivative work must be distributed under the same license as the original. Full terms at https://creativecommons.org/licenses/by-nc-sa/4.0/.</p> |

| <u>Source name</u> | <u>Last accessed</u> | <u>Source URL</u> |
|--------------------|----------------------|-------------------|
|--------------------|----------------------|-------------------|

2007 Population
Census of Peru, via
IPUMS

06-Apr-21 <https://international.ipums.org/>

OCHA administrative
boundary and P-code
data via HDX

6-Apr-21 <https://data.humdata.org/dataset/limites-de-peru>

| <u>Field heading</u> | <u>HXL code</u> | <u>Description</u> | <u>Sources</u> |
|---------------------------|--------------------------------------|---|---|
| admin2_name | #adm2+name | Name of administrative unit | OCHA administrative boundary and pcode data |
| admin2_pcode | #adm2+code | Unique geographic identification code | OCHA administrative boundary and pcode data |
| admin1_name | #adm1+name | Name of administrative unit | OCHA administrative boundary and pcode data |
| admin1_pcode | #adm1+code | Unique geographic identification code | OCHA administrative boundary and pcode data |
| admin0_name | #country+name | Name of administrative unit | OCHA administrative boundary and pcode data |
| admin0_pcode | #country+code | Unique geographic identification code | OCHA administrative boundary and pcode data |
| number_of_named_languages | #indicator+lang+num | Number of named (non Other) languages reported in area (excluding those with less than 0.1 coverage in this area) | Calculated field |
| main_language | #indicator+lang+main | Name of reported language with highest attributed population in area | Calculated field |
| main_language_share | #indicator+lang+main+pct | Percentage share of population in area attributed to main language | Calculated field |
| .. | .. | .. | .. |
| Language fields | #indicator+lang+pct+iso639_3_ABC | communication in their households | 2007 census data |
| Other language | #indicator+lang+pct+iso639_3_unknown | communication in their households is not a named language | 2007 census data |
| .. | .. | .. | .. |

| | | | |
|--------------------|---|---|------------------|
| speaks_eng_ | #indicator+lang+understand+pct+iso639 3_ABC | Percentage of population aged 15 and over who can speak english | 2007 census data |
| illiteracy_ | #population+total+pct+illiterate+age15up | Percentage of population aged 15 and over who are illiterate | 2007 census data |
| illiteracy__male | #population+m+pct+illiterate+age15up | Percentage of male population aged 15 and over who are illiterate | 2007 census data |
| illiteracy__female | #population+f+pct+illiterate+age15up | Percentage of female population aged 15 and over who are illiterate | 2007 census data |
| pop_total | #population+total | Number of total population | 2007 census data |
| pop_male | #population+m | Number of male population | 2007 census data |
| pop_female | #population+f | Number of female population | 2007 census data |
| language_data_yn | #meta+data+bool | record | |
| data_confidence | #meta+confidence | Assessment of data quality (high/medium/low) | |
| notes | #meta+text | Notes and clarifications on data | |

Data confidence desc| Examples

High Data from well regarded government census or other specialist organisation.

Medium Data from quality source but without clear methodology, or inverse.

Low Data known to be erroneous, or based on non statistical or imprecise data.

Factors increasing con Data is corroborated. Data can be directly applied to administrative areas without aggregation or manipulation.

Factors decreasing cor Data is disputed. Data requires transformation that may affect accuracy, for example combining admin 2 data to form admin 1 data.

| <u>Language</u> | <u>ISO 639 3 code</u> | <u>Alternative names and spellings</u> |
|------------------------|------------------------------|---|
| Asháninka | cni | Ashinanca |
| Aymará | ayr | |
| Quechua | que | Kichua |
| Spanish | spa | Español |