

## Metadata

<b>Version</b>	1
<b>Created by</b>	Translators Without Borders
<b>Uploaded on</b>	mar/20
<b>Notes and caveats</b>	<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>The methodology by which literacy was measured during the census is not known.</p> <p>Language and dialect naming conventions in the census are sometimes unclear and in some cases have been grouped or allocated to the Other Languages field.</p> <p>Language shares below 0.1% of an area population have been discarded. As a result languages may not be present at Medium administrative levels.</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> <p>Empty values represent non-existent data and should not be treated as zero values.</p> <p>Data on languages spoken and literacy were only available for the year 2000 in the admin 0, admin1 and admin2 levels. Some provinces and districts were created since, so there is no data available for some entries.</p> <p>Data on languages spoken made available publicly do not state which languages the answer "Other" refers to. In cases where Thai language literacy is below 90%, a note was included with the language to which it probably refers based on Wikipedia's Demographic of Thailand page.</p> <p>Spoken language data in the 2000 Thailand Population Census was measured by asking if the primary language spoken at home was (1) Only Thai, (2) Thai and other or (3) Only other. The data on specific "other" spoken languages is available exclusively at admin level 0 since it comes from the 2000 Census Reports, where it is described in text format.</p>
<b>Copyright and terms of use</b>	<p>© Translators without Borders 2020</p> <p>Data is available under a Attribution-NonCommercial-ShareAlike 4.0 International license (CC BY-NC-SA 4.0). 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 <a href="https://creativecommons.org/licenses/by-nc-sa/4.0/">https://creativecommons.org/licenses/by-nc-sa/4.0/</a>.</p>

<u>Source name</u>	<u>Last accessed</u>	<u>Source URL</u>
OCHA administrative boundary and P-code data via HDX (version at Mar 15 2020)	15/mar/20	<a href="https://data.humdata.org/dataset/thailand-administrative-boundaries">https://data.humdata.org/dataset/thailand-administrative-boundaries</a>
National Statistical Office (Thailand). Thailand Population and Housing Census 2010.	15/mar/20	<a href="http://service.nso.go.th/nso/nso_center/project/search_center/23project-en.htm">http://service.nso.go.th/nso/nso_center/project/search_center/23project-en.htm</a>
National Statistical Office (Thailand). Thailand Population and Housing Census 2000.	15/mar/20	<a href="http://web.nso.go.th/en/census/poph/cen_poph.htm">http://web.nso.go.th/en/census/poph/cen_poph.htm</a>
Demographics of Thailand	15/mar/20	<a href="https://en.wikipedia.org/wiki/Demographics_of_Thailand">https://en.wikipedia.org/wiki/Demographics_of_Thailand</a>

<b>Field heading</b>	<b>HXL code</b>	<b>Description</b>	<b>Sources</b>
admin2_name	#adm2+name	Name of administrative unit	OCHA administrative boundary and P-code data
admin2_pcode	#adm2+code	Unique geographic identification code	OCHA administrative boundary and P-code data
admin1_name	#adm1+name	Name of administrative unit	OCHA administrative boundary and P-code data
admin1_pcode	#adm1+code	Unique geographic identification code	OCHA administrative boundary and P-code data
admin0_name	#country+name	Name of administrative unit	OCHA administrative boundary and P-code data
admin0_pcode	#country+code	Unique geographic identification code	OCHA administrative boundary and P-code data
number_of_name{	#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 Mediumest attributed population in area	Calculated field
main_language_sh	#indicator+lang+main+pct	Percentage share of population in area attributed to main language	Calculated field
..	..	..	..
Language fields	#indicator+lang+pct+iso639-3_ABC	Percentage of population who generally speak this language at home	National Statistical Office (Thailand). Thailand Population and Housing Census 2000.
Other language	#indicator+lang+pct+iso639-3_unknown	Percentage of population who generally speak a non-named language at home	National Statistical Office (Thailand). Thailand Population and Housing Census 2000.
..	..	..	..
pop_total	#population+total	Number of total population in 2010	National Statistical Office (Thailand). Thailand Population and Housing Census 2010.
pop_male	#population+m+total	Number of male population in 2010	National Statistical Office (Thailand). Thailand Population and Housing Census 2010.
pop_female	#population+f+total	Number of female population in 2010	National Statistical Office (Thailand). Thailand Population and Housing Census 2010.
literacy_all	#population+total+pct+literate+age10up	Percentage of population aged 10 and over who are literate	National Statistical Office (Thailand). Thailand Population and Housing Census 2000.
literacy_male	#population+m+pct+literate+age10up	Percentage of male population aged 10 and over who are literate	National Statistical Office (Thailand). Thailand Population and Housing Census 2000.
literacy_female	#population+f+pct+literate+age10up	Percentage of female population aged 10 and over who are literate	National Statistical Office (Thailand). Thailand Population and Housing Census 2000.

language_data_yn	#meta+data+bool	Binary field describing whether language data is available in this
data_confidence	#meta+confidence	Assessment of data quality (Medium/medium/low)
notes	#meta+text	Notes and clarifications

### **Data confidence**

<b><u>descriptions</u></b>	<b><u>Examples</u></b>
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 confidence	Data is corroborated. Data can be directly applied to administrative areas without aggregation or manipulation.
Factors decreasing confidence	Data is disputed. Data requires transformation that may affect accuracy for example combining admin 2 data to form admin 1 data.

<b><u>Language</u></b>	<b><u>ISO 639-3 code</u></b>	<b><u>Alternative names and spellings</u></b>
Thai	tha	
Maylay-yawi	mfa	Kelantan-Pattani Malay
Chinese	zho	Mandarin
Karen	ksw	Sgaw Karen or Sgaw Kayin
Thaikueng	unknown	
Morn	mnw	Mon
Lao-krung	krng	Brao Krung
Hmong/Mea	hmn	
Local languages	unknown	
Burmese	mya	Burmeae
Lao	tts	Isan
Khmer	khm	Cambodian
Japanese	jpn	
English	eng	