

Metadata

Version	1
Created by	Translators without Borders and University College London
Uploaded on	Feb-20
Notes and caveats	<p>Mother tongue is defined in the Pakistan 1998 census as the language which parents speak with their children in daily life.</p> <p>No language data is available for Saraiki at admin 2 level. Saraiki speakers at this level may be represented in Other, or in Punjab, which Saraiki is sometimes treated as a dialect of.</p> <p>Language shares below 0.1% of an area population have been discarded.</p> <p>All decimal values have been rounded to a maximum of 3 decimal places.</p> <p>Though the census covered all of Pakistan, Azad Kashmir and Gilgit-Baltistan, the sample provided by IPUMS does not include the Federally Administered Tribal Areas, Azad Kashmir, and Gilgit-Baltistan provinces. Additionally, the administrative boundaries provided by OCHA did not include the former Kashmir territories on either side of the Line of Control. Future versions of this dataset could include language information for these regions if the data becomes accessible.</p>
Copyright and terms of use	<p>© Translators without Borders 2020</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>

Source name	Last accessed	Source URL
OCHA FISS administrative level 0-3 population statistics	26-Jul-19	https://data.humdata.org/dataset/pakistan-administrative-level-0-3-population-statistics
Pakistan Bureau of Statistics 1998 census data	27-Jul-19	http://www.pbs.gov.pk/content/population-mother-tongue
IPUMS 1998 Pakistan census microdata	28-Jul-19	https://international.ipums.org/international/
Pakistan Social and Living Standards Measurement Survey 2014-15	29-Jul-19	http://www.pbs.gov.pk/sites/default/files//pslm/publications/PSLM_2014-15_National-Provincial-District_report.pdf
Multiple Indicator Cluster Survey, 2009	30-Jul-19	http://fata.gov.pk/cp/uploads/downloads/MICS.pdf
Kech 2011 District Education Strategy	31-Jul-19	http://emis.gob.pk/Uploads/KECH%20(TURBAT)%20DISTRICT%20EDUCATION%20PLAN%20FOR%202016-2017%20TO%202020-2021.pdf
Panjgur 2013 District Education Strategy	01-Aug-19	http://emis.gob.pk/Uploads/PANJGUR%20DISTRICT%20EDUCATION%20PLAN%20FOR%202016-2017%20TO%202020-2021.pdf
North Waziristan 1998 census data	02-Aug-19	https://pdfs.semanticscholar.org/2949/d346fcdea580d286d02736a8be8b2fc8c709.pdf

Field heading	HXL code	Description	Sources
admin2_name	#adm2+name	Name of administrative unit	OCHA FISS administrative level 0-3 population statistics, original source Pakistan Bureau of Statistics
admin2_pcode	#adm2+code	Unique geographic identification code	
admin1_name	#adm1+name	Name of administrative unit	
admin1_pcode	#adm1+code	Unique geographic identification code	
admin0_name	#country+name	Name of administrative unit	
admin0_pcode	#country+code	Unique geographic identification code	
main_language	#indicator+lang+main	Name of the mother tongue language with highest usage in this area	Calculated field
main_language_share	#indicator+lang+main+pct	Percentage share of the dominant mother tongue language	
Balochi	#indicator+lang+pct+is639-3_bal	Percentage of population whose mother tongue is Balochi	Admin 0-1: Pakistan Bureau of Statistics 1998 census data. Admin 2: IPUMS 1998 Pakistan census microdata.
Punjabi	#indicator+lang+pct+is639-3_pan	Percentage of population whose mother tongue is Punjabi	
Pushto	#indicator+lang+pct+is639-3_pus	Percentage of population whose mother tongue is Pushto	
Sindhi	#indicator+lang+pct+is639-3_snd	Percentage of population whose mother tongue is Sindhi	
Urdu	#indicator+lang+pct+is639-3_urd	Percentage of population whose mother tongue is Urdu	
Saraiki	#indicator+lang+pct+is639-3_skr	Percentage of population whose mother tongue is Saraiki	
Other	#indicator+lang+pct+is639-3_unknown	Percentage of population whose mother tongue is Other	
pop_total	#population+total	Number of total population	OCHA FISS administrative level 0-3 population statistics, original source Pakistan Bureau of Statistics
pop_urban_%	#population+urban+pct	Percentage of total population residing in urban areas	
pop_male	#population+m	Number of male population	
pop_female	#population+f	Number of female population	
pop_transgender	#population+i	Number of transgender population	
literacy_%	#population+total+pct+literate+age_gt_10	Percentage of population above 10 who are literate	Mostly Pakistan Social and Living Standards Measurement Survey 2014-15, Pakistan Bureau of Statistics. Updates and corrections made to reflect changes to administrative structure since publication (see note field on individual records). FATA admin2 literacy data from 2009 Multiple Indicator Cluster Survey, Government of Pakistan. Kech and Panjgur admin2 data from 2011 and 2013 District Education Strategies. North Waziristan admin2 data from 1998 census.
literacy_%_male	#population+m+pct+literate+age_gt_10	Percentage of male population above 10 who are literate	
literacy_%_female	#population+f+pct+literate+age_gt_10	Percentage of female population above 10 who are literate	
literacy_%_urban	#population+urban+pct+literate+age_gt_10	Percentage of urban population above 10 who are literate	Pakistan Social and Living Standards Measurement Survey 2014-15, Pakistan Bureau of Statistics. Updates and corrections made to reflect changes
literacy_%_rural	#population+rural+pct+literate+age_gt_10	Percentage of rural population above 10 who are literate	

			to administrative structure since publication (see note field on individual records).
language_data_yn	#meta+data+bool	Binary field describing whether language data is available in this record	-
data_confidence	#meta+confidence	Assessment of data quality (high/medium/low)	-
notes	#meta+text	Notes and clarifications on data	-

<u>Data confidence descriptions</u>	<u>Examples</u>
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.

<u>Language</u>	<u>ISO 639-3 code</u>	<u>Alternative names and spellings</u>
Balochi	bal	
Punjabi	pan	
Pushto	pus	Pashto
Sindhi	snd	
Urdu	urd	
Saraiki	skr	Seraiki