

PANTHEON: A DATASET OF GLOBALLY FAMOUS BIOGRAPHIES

Pantheon is a manually verified dataset of famous people that was created by Yu et al. [1]. It includes 11,341 biographies that have presence in more than 25 languages in Wikipedia by May 2013.

Included attributes:

name	– First- and surname of the person corresponding to the biography
numlangs	– Number of language editions in which this biography is available on Wikipedia
birthcity	– City of birth connected to the person
birthstate	– US state corresponding to the birthcity (if applicable)
countryName	– Name of the country corresponding to the birthcity
countryCode	– Two-digit country code corresponding to the birthcity
countryCode3	– Three-digit country code corresponding to the birthcity
LAT/LON	– Latitude and longitude corresponding to the birthcity
continentName	– Continent corresponding to the birthcity
birthyear	– Year of birth of the notable person
gender	– Gender of the person
occupation	– Occupation of the person
industry/domain	– Industry and domain related to the occupation of the person
TotalPageViews	– Total of Wikipedia page views on the biography
PageViewsEnglish	– Total of page views on English Wikipedia pages of the biography
PageViewsNonEnglish	– Total of page views on non-english Wikipedia pages of the biography
L_star	– Effective number of languages (see more information at [1])
HPI	– Historical Popularity Index (see more information at [1])

Further information:

[1] Yu, A. Z., et al. (2016). Pantheon 1.0, a manually verified dataset of globally famous biographies. *Scientific Data* 2:150075. doi: 10.1038/sdata.2015.75