**Onomap2**

Onomap2 is based upon the global Worldnames2 database of 1.2+ billion names, described in O’Brien and Longley (O'Brien & Longley, 2018). It estimates the most probable country of origin using a system of weighted look-up tables. It is designed for the widest range of international applications.

There are ambiguities inherent in this process, arising from promulgation of Old World English, Spanish or Portuguese names in the Americas and Australasia. In other cases, such as the Balkans, former Soviet Union and parts of Central and South America, discrete national boundaries may be super-imposed upon more continuous distributions of family groups. In still other settings, such as some Middle Eastern countries, resident populations may be dominated by guest workers originating from elsewhere in the world. For these reasons we provide estimates of the top three countries for each name.

We provide two different measures. We estimate the relative frequency of name occurrences in every country and record the relative frequencies of occurrence in the three highest countries – the relative frequencies across all countries sum to unity. Additionally, absolute counts are recorded for the country in which the largest estimated number of bearers of the name resides.

**Reference:**

[O'Brien, O., & Longley, P. (2018). Given and family names as global spatial data infrastructure. In P. Longley, J. Cheshire, & A. Singleton (Eds.),](https://www.uclpress.co.uk/products/114094) *[Consumer Data Research](https://www.uclpress.co.uk/products/114094)* [(pp. 53-67). London: UCL Press.](https://www.uclpress.co.uk/products/114094)