



GRC Database Information
Nieuwe Prinsengracht 80-hs
1018 VV AMSTERDAM
The Netherlands
020-528 61 20 (telephone)
020-422 93 11 (fax)
graham@grcdi.nl
<http://www.grcdi.nl>

Given names

NB: Tables are provided as is. The publisher is not responsible or liable for errors or damage resulting from the use of these tables.

Highlighted text indicates additions/alterations since the last document version.

The attributes of the table are as follows:

- Records contain the given name as they appeared in the incoming record.
- The file contains for some names a corrected version. Corrections are only made when there is a clear indication that a name is incorrect and the correct version is obvious. Thus it is never correct to assume that *Boby* should be *Bobby*, or that *Bob* should be *Robert*. However, *Andr* can clearly be corrected to *André*.
- This data comes from real World sources, and can contain real World data errors. Though we mark data which we know is not correct, there can be no guarantee that a combination is correct on the ground.
- We concentrate resources on data found most often and most recently. Data found more often is likely to be more accurate.
- Only records containing names found five times or more in real world data are released.

The data is provided in Windows code page 1252, complying to ISO-8859-1, MS-DOS code page 850. Data is in Latin script only.

Table structure

<i>Field name</i>	<i>Field type</i>	<i>Field length</i>	<i>Contents</i>
URN	Character	10	A unique code for the combination of data in this field. This number should be used as a reference if there are any queries about the data.
ISO2	Character	3	The ISO 3166* 2-digit code for this country.
ISO3	Character	3	The ISO 3166* 3-digit code for this

			country.
ISONUMERIC	Character	3	The ISO 3166* 3-digit numeric code for this country.
NAME	Character	70	The given name as it appeared in the incoming file.
CORRECTED	Character	70	A corrected version of the NAME string where errors are clear and correctable
GENDER	Character	70	A gender indicator: M =Male, F =Female, empty =unknown
COUNT	Numeric	10	The number of times that the NAME has been found in real world data analyzed by GRC Database Information.
INITIALS	Logical	1	The data in the NAME field are initials. For example: <i>A.J.P.</i>
INCL_INIT	Logical	1	The data in the NAME field includes initials as well as other name string(s). For example: <i>Edward L.P.</i>
INCL_SALUT	Logical	1	The data in the NAME field contains a form of address or a seniority indicator. For example: <i>Mr William, Fred Jr.</i>
MULTIPLE	Logical	1	The data in the NAME field contains multiple given names. For example: <i>Claire & Robert, Jean et Thérèse</i>
DATE_ADDED	Date	8	Date that the record was added.
DATE_CHANG	Date	8	Date that the given name string was last encountered in a file. Note: this does NOT refer to any changes which are made in any other fields in the file, which are constantly being updated.
LINKEDTO	Character	254	Other names with which this name may be related. These names are delimited with ‘ ’ (excluding the quotation marks). For example: JIM JAMES JAMIE JIMMY

We advise extreme caution is attempting any genderisation process on the basis of given names, but as this data is often used for this process, we have released a second table, containing gender information distilled from the main file. This table (called GENDER) has data in this format:

ISO2	Character	3	The ISO 3166-2 code for this country. Genderisation must be attempted on a country-by-country basis rather than globally, as the same name strings are used for different genders in different places.
------	-----------	---	--

NAME	Character	70	The UPPER CASE version of the name as found in the main file
MALE	Numeric	10	The number of times that this name has been found to indicate a MALE for this country
MALE_PERC	Numeric	8,2	The percentage of all gendered occurrences of this name found to be male for this country
FEMALE	Numeric	10	The number of times that this name has been found to indicate a FEMALE for this country
FEMALE_PER	Numeric	8,2	The percentage of all gendered occurrences of this name found to be female for this country
UNIVERSE	Numeric	10	The total number of gendered occurrences found in the main file for this country (i.e. MALE+FEMALE)

End

Graham Rhind
GRC Database Information
Nieuwe Prinsengracht 80-hs
1018 VV AMSTERDAM
020-528 61 20 (telephone)
020-422 93 11 ('fax)
graham@grcdi.nl
<http://www.grcdi.nl>