

pdNickname – Getting Started Guide

pdNickname is a unique nearly 50,000 record database designed to facilitate comparing sets of first name data based on nicknames, diminutives, pet names, variations and given names.

Matching and merging names can be tricky. How do you relate William Smith with Bill Smith? The *pdNickname* database can be utilized to match names that are dissimilar because one has a given first name while another has a nickname or other variation.

This one-of-a-kind package is particularly suited for matching American and Americanized first names, nicknames, diminutives, pet names and variants as well as ethnic first names and variations common in the United States.

Included Database Files

Four different file formats are included to insure compatibility with any database system. Each of the files contains the same data.

- pdNickname.xls (MS Excel)
- pdNickname.csv (comma-delimited ASCII text—CSV)
- pdNickname.txt (fixed length ASCII text)
- pdNickname.dbf (dBase/FoxPro—easily imports into MS Access)

Included User Documentation

- README.txt (general information)
- This Getting Started Guide
- pdNickname-Layout.pdf (layout & data definitions)
- pdNickname-Sample-Program.prg (sample program for comparing first names)
- Site License (see Licensing below)

Version Number

The version number of each copy of *pdNickname* is written in the first or second row of the first column of the database files in **X.X.X** format. The first numbers is the version of the release. The number after the first dot is the update for the version indicated. The number after the second dot references a minor revision.

Layout & Data Definitions

File specifications can be found in the included *pdNickname-Layout.pdf* file. This gives the name of each field, the character type (character/alphanumeric or numeric), the width (including any decimal points, the starting point and the ending point) and the definition of the field contents.

Instructions

Users can match against the **NAME** and **VARIATION** fields to compare the names on their lists with related nicknames, diminutives, pet names, variations and given names.

While there are numerous ways to work with *pdNickname* data, when two first names are being evaluated, the user typically scans all occurrence of first name #1 in the *pdNickname* **NAME** field, and simultaneously compares the names in the *pdNickname* **VARIATION** field with first name #2 to determine if there is a match.

The relationship between the names in the *pdNickname* **NAME** and **VARIATION** fields are flagged in **RELFLAG** field as either “close” or “more distant”. Additional accuracy can be achieved if only close relationships are considered. However, with residential lists, the margin of error generally does not substantially increase if the more distant relationships are included. Conversely, it is usually recommended that only close relationships be considered when processing business and organizational lists.

Matching is against the names in the list or lists being processed. This is relatively straightforward if the first names are parsed from remainder of the name and in a separate field—which is the recommended technique for name processing. Nevertheless, matching can often be achieved even when the names are unparsed. In these cases list names should be compared to trimmed versions of the names in the *pdNickname* database (i.e., with trailing blank spaces removed). This will not work, however, if the name is preceded with a title, such as Mr., Mrs., Ms, Rev. or Dr.

The names in the *pdNickname* database are in UPPER CASE. Therefore, it is important that the list names being matched against are converted to UPPER

CASE as well. Normally such conversions are performed in the matching program and do not affect the formatting in the actual list.

Gender is included as a field in the *pdNickname* database; however, it should not be considered a comprehensive gender coding database because first names without a nickname or other variation are not included.

Sample Program

A sample program showing a technique for comparing first names can be found in the included *pdNickname-Sample-Program.prg* file. It is a function that accepts two first names and returns a value indicating if and how the names are related. The program is written in Visual FoxPro, but the concepts are translatable to almost any programming language.

Licensing

If purchased as a standalone product, the Site License for this product can be found in the included *pdNickname-Site-License.pdf* file. If purchased as part of *pdSuite Master Collection*, it can be found in the included *pdSuiteMC-Site-License.pdf* file. Users are allowed to install *pdCensus2010* on all computers in the same building within a single company or organization (see the license for details). A Developer License is also available for separate purchase granting users the right to incorporate the information in for-profit services and products.

pdNickname - Name & Nickname Database

Users match against the NAME fields to compare the names on their lists with related nicknames, diminutives, pet names, variations, and given names in the VARIATION field.

COLUMN HEADINGS

- FIELD: name of the field.
- TYPE: field type; Chr = character or alphanumeric, Num = numeric.
- WIDTH: field width.
- DEC: number of decimal points.
- FROM: field starting point.
- TO: field ending point.
- DESCRIPTION: data definition of the field contents.

Field	Type	Width	Dec	Start	End	Description
PEACOCK_ID	Chr	10		1	10	Unique ID number assigned to each record
NAME	Chr	25		11	35	First name
VARIATION	Chr	25		36	60	Nickname, diminutive, pet name, variation, or given name
RELFLAG	Chr	1		61	61	Relationship Flag (key): 1 = Close relationship between the name and variation (common variants) 2 = More distant relationship between the name and variation (less common variants)
VARCNT	Num	10	0	62	71	Number of variations of the name
VARNUM	Num	10	0	72	81	Sequential variation number
GENDER	Chr	1		82	82	Gender (key): M = Male F = Female U = Undetermined

January 1, 2013

pdNickname - Site License

- I. This Site License grants to the Licensee the right to install the licensed version of **pdNickname** (hereinafter, 'information') on all computers in the same building within a single company or organization. Separate Site Licenses must be purchased for each building the information is used in.
- II. The information may only be used by the employees of the Licensee. If the Licensee is an educational institution, the data may only be used by enrolled students, faculty, teaching assistants, and administrators.
- III. Temporary employees, contractors, and consultants of the Licensee who work on-site at the Licensee's facility may also use the information in connection with the operation of the business of the Licensee. Any copies of the information used by temporary employees, contractors, and consultants must be removed from such individual's computers once they cease working at the Licensee's facility.
- IV. The information cannot be used to provide services or products to customers or other third parties, whether for-profit or given away. A Developer License must be purchased separately by the Licensee to incorporate the information in for-profit services and products.
- V. The Licensee is required to use commercially reasonable efforts to protect the information and restrict network or any other access to the information by anyone inside or outside of the Licensee's facility who is not authorized to use the information.
- VI. The Licensee owns the media on which the information is recorded or fixed, but the Licensee acknowledges that Peacock Data, Inc. and its licensors retain ownership of the information itself.
- VII. The Licensee may not transfer or assign its rights under this license to another party without Peacock Data, Inc.'s prior written consent.

- VIII. Peacock Data, Inc. may revoke the rights granted by this license upon a violation of any provision herein by the Licensee.
- IX. This Site License is governed by Peacock Data, Inc.'s Terms of Service and Privacy Policy, and the laws and regulations of the United States and the State of California.

Liability Notice

PEACOCK DATA, INC. PROVIDES DATABASE SERVICES AND PRODUCTS FOR LAWFUL PURPOSES ONLY AND EXPRESSLY FORBIDS THEIR USE IN ANY UNLAWFUL MANNER. FURTHERMORE, PEACOCK DATA, INC. SHALL NOT BE LIABLE TO CUSTOMERS FOR THE COMPLETE ACCURACY OF ALL DATA EXAMINED AND/OR ALTERED, NOR FOR THE COMPLETE ACCURACY OF ALL PRODUCTS LICENSED, NOR FOR ANY LOSS OR INJURY ARISING OUT OF OR CAUSED, IN WHOLE OR IN PART, BY PEACOCK DATA, INC.'S NEGLIGENT ACTS OR OMISSIONS IN PROCURING, COMPILING, COLLECTING, PROCESSING, DEVELOPING, COMMUNICATING OR DELIVERING THE DATA AND/OR PRODUCTS MENTIONED HEREIN, DIRECTLY OR INDIRECTLY TO THE CUSTOMER. PEACOCK DATA, INC.'S SOLE LIABILITY AND THE CUSTOMER'S SOLE REMEDY, REGARDLESS OF THE FORM OF LEGAL ACTION TAKEN BY THE CUSTOMER, WHETHER IN TORT OR CONTRACT, SHALL NOT EXCEED THE REFUND OF THE AMOUNTS, IF ANY, PAID BY THE CUSTOMER. CUSTOMERS AGREE PEACOCK DATA, INC. SHALL NOT BE LIABLE FOR LOST PROFITS OR ANY INDIRECT, INCIDENTAL, SPECIAL OR CONSEQUENTIAL DAMAGES, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Returns and Refunds

The files on disks, tapes, and media retrieved via the Internet will be replaced at no additional charge if the file(s) and/or media is/are found to be corrupted or otherwise physically defective, if reported to Peacock Data, Inc. within 30 days of purchase. Because database products are non-tangible, there are no refunds under any circumstances. Purchases are governed by Peacock Data, Inc.'s Terms of Service and Privacy Policy.

Copyright Notice

pdNickname is Copyright © 2009 Peacock Data, Inc. All Right Reserved.