

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 *pdNickname* on all computers 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.