



DATA DIRECT
A U S T R A L I A

31-33 Chifley Drive Preston VIC 3072
PO Box 2096 Preston VIC 3072
Phone + 61 3 9484 3282
Fax + 61 3 9480 0082
dda@datadirect.com.au
www.datadirect.com.au

Cleaning & Updating Your Database



Database Services

Rather than utilise generic third party off the shelf data manipulation and repair packages, Data Direct Australia has called upon the collective knowledge, spanning some thirty five years of its highly experienced programmers and database managers to develop its own suite of highly effective and intuitive data repair and standardisation programs. These programs have been designed, developed and continually enhanced by the people who use them on a daily basis, hence they understand the nuts and bolts and the often messy realities of the real life situations that arise in many databases.

The array of programs that we have available at our disposal includes, but is not limited to the following:

Punctuation Removal

To comply with Australia Post's "Correct Addressing" Guidelines to attract maximum postal discounts.

Floating Field Standardisation

To facilitate the standardisation of "like" data into fixed constant fields to assist in data segmentation, analysis and deduplication functions.

Name Line Split

To segment a "Contact Name" field into separate Title, First Name & Surname fields utilising look-up exception tables to allow for multi word names such as "Van Der Har", "De Vincenzo", "Van Merlo" etc.

First Word / Last Word Extraction

A basic facility to be able to extract either the first word or last word in any data field.

Suburb Line Split

If not supplied individually, a facility to split a suburb line into separate Suburb, State & Postcode fields.

Field Concatenation

The ability to join contents of different fields together to form new fields (e.g.; Title, Firstname & Surname to form Nameline)



Upper Casing

A facility to totally upper case either individually selected or all fields in a database. Particularly useful if mixed or inconsistent casing is supplied in raw data.

Proper Casing (Utilising exceptions database)

Our Proper Casing facility not only upper cases the first letter of each word but also reads in an exception database which caters for hyphenated names, Apostrophe's (e.g.; O'Connel), Mac's and Mc's as well as abbreviated company names such as "BHP", "AMP", "RACV" etc.

Suburb, State & Postcode Allocation / Correction

The ability to populate missing suburbs where postcode is available & the postcode only has one related suburb. The ability to populate missing postcodes where suburb & state are available. The ability to allocate states where postcode is available. The ability to interrogate, validate & offer repair suggestions & execute them where either suburb, state or postcodes are found to be missing or incorrect according to current Australia Post lookup tables.

Personal Title Allocation based on Christian Names

Based on a lookup table of Christian Names formulated over the years, we are able to allocate either "Mr" for Males or "Ms" for females where the gender associated with a Christian name is not considered to be ambiguous. In other words the Christian name as a rule is not used across both genders. Names such as "Chris", "Pat", "Robin" etc are not allocated a title as they are deemed to be "Cross Gender".

Personal Title Gender Validation

Based on the same table as used for title allocation, we are able to validate & repair existing titles in a database to identify invalid allocations such as "Miss George Smith" or "Mr Sally Adams" etc.

Street Type Standardisation (Abbreviated)

This facility allows us to interrogate street types in address fields and convert them from their extended format to an abbreviated format to facilitate consistency which is particularly useful for deduplication and data matching / linking purposes.

Street Type Standardisation (Extended)

This is the reverse of the previous process which is sometimes used after a deduplication process is executed due to the client preferring an extended format for aesthetics.