

Using the 'Kit-Form' Database

'Kit-Form' Database Concept

All relational databases are made up of tables. If such a database is broken down into its component tables, the result is a set of objects that have little or no value as discrete entities. However, these objects may be regarded as a kit of parts that may be assembled to produce a fully functional database. This is the fundamental concept behind the 'kit-form' database.

The kit consists of a series of files. Each file contains a table from the database as a totally portable ASCII file. The result is a dump of the database that may be easily incorporated into any relational database management system on any platform.

Files and Formats

The files of the 'kit-form' database may be found in the DBKIT directory. The names of the files generally conform to the name of the table in the database to which they correspond. The only exceptions are a small number of tables whose names were longer than 8 bytes. These have been truncated to give DOS-compatible filenames but may be unambiguously associated with table names in the database structural documentation. All files have the file extension CSV.

The file format is very simple and straightforward. The files are in standard ASCII code. Note that as the CD-ROM was assembled on a PC, the record terminators are 2-byte CRLF sequences, not the single-byte terminators found on UNIX systems.

The first record of each file contains the table field names separated by commas. The second and subsequent records contain the table data with one row per record. Obviously, the order of the fields in the data records corresponds to the order of the field names in the header.

In general, text fields have no delimiters except for the leading and trailing commas. However, the text fields from some tables included embedded commas. In these cases, the text fields are enclosed by double quote (") symbols.

This file format is sometimes known as 'comma separated value' or 'CSV' format.

Suggested Applications for the 'Kit-Form' Database

The 'kit-form' concept was conceived to provide a totally portable export mechanism for BODC's project databases. The 'target user' was someone who wished to import data in bulk into their own database, either preserving the BODC structures or reformatting some or all of the data into an alternative schema. This has been done with considerable success by a number of users from 'kit-form' databases on previous BODC electronic publications.

However, this is not the only way in which the 'kit-form' database files may be used. Part of the software interfaces for the North Sea Project and BOFS CD-ROMs actually used the 'kit-form' database files. It is therefore perfectly possible for users who do not have access to relational database management systems to develop conventional application programs against 'kit-form' database files. They are, after all, just a set of very ordinary data files.

One final suggestion. Spreadsheet users might find it useful to load up the EVENT.CSV file. This is an inventory of the data collection events that make up the database. Most modern spreadsheet packages have limited database query capabilities. If these are applied to the data from EVENT, users can get a feel of properties of the data such as 'what came from where' and therefore determine whether the data set is of interest.