



**TOGETHER**  
*for a sustainable future*

## OCCASION

This publication has been made available to the public on the occasion of the 50<sup>th</sup> anniversary of the United Nations Industrial Development Organisation.



**TOGETHER**  
*for a sustainable future*

## DISCLAIMER

This document has been produced without formal United Nations editing. The designations employed and the presentation of the material in this document do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations Industrial Development Organization (UNIDO) concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries, or its economic system or degree of development. Designations such as “developed”, “industrialized” and “developing” are intended for statistical convenience and do not necessarily express a judgment about the stage reached by a particular country or area in the development process. Mention of firm names or commercial products does not constitute an endorsement by UNIDO.

## FAIR USE POLICY

Any part of this publication may be quoted and referenced for educational and research purposes without additional permission from UNIDO. However, those who make use of quoting and referencing this publication are requested to follow the Fair Use Policy of giving due credit to UNIDO.

## CONTACT

Please contact [publications@unido.org](mailto:publications@unido.org) for further information concerning UNIDO publications.

For more information about UNIDO, please visit us at [www.unido.org](http://www.unido.org)

21195

**Final Report on  
the  
Development of  
Statistical Data Base Query for Windows**

**prepared by  
Hungarian Central Statistical Office  
and  
SOLTRONIC B.V.**

**July 1995**

## **Introduction**

Statistical Data Base Query for Windows (abbreviated to SDBQ for Windows) is the software that transforms the way you use large, even multimillion records databases on CD-ROM. This version is the next generation of the software, named Statistical Data Base Query, (SDBQ) developed for DOS from 1991 to 1993. The best standardized media is the Compact Disc-Read Only Memory (CD-ROM) therefore it is an ideal mass memory for the distribution and the use of information which are useful for the users of the government or other private sectors. The statistical user community welcomed the new way of publication of the statistical data in Hungary but also required an enhanced and improved version of the software. After it was distributed for the public the questionnaires returned from the users by many statistical offices and research institutes were evaluated and processed.

## **Hardware and software requirements**

The software runs on 386SX or higher PC/AT compatible microprocessor based personal computer with minimum of 4 Mbytes of RAM and it also expects approximately 20 Mbytes of free hard disk space. The use of 486 processor based PCs with double speed or quad speed CD-ROM drives are recommended. The software itself is one complete Win16 program which runs under Windows operating system version 3.1. The PC should be hooked up with a CD-ROM drive and the using Microsoft CD-ROM Extension interface (driver and program) is needed. It is also a must that the databases are prepared in proprietary format of the HCSO developed for unit-record databases. The content of the CD-ROM can be any statistical database, population, vital statistics or industrial data.

The software allows the user to select a database, to display the content and prepare tables for immediate study, presentation or publication. It makes possible for the user to compose, edit, modify save layout of

statistical tables and manage as it required. The mapping module allows the user to map the data, on geographical units using the GIS data provided by the HCSO, if it is relevant.

The input files, either the statistical data or the map (Geographical Information System) data, are specially designed for the fast CD-ROM access method.

### **User interface**

The software uses the graphical tools for the information processing and communication between the computer and the user and passes the features of the Windows to the user. It has clear communication interface: provides the Window based menu system and communication. As usual, it sends messages and accepts user inputs via dialog boxes or text boxes, displays information in list boxes, lets the user to make confirmation using buttons or check boxes. On-line help system is also available for getting information on how to use the SDBQ for Windows. This product is also bilingual as its predecessors were implemented under DOS.

### **Functionality**

This system has the same or improved functionality as the SDBQ for DOS version has. Despite of the efforts to create the same functionality, menu and sub-menus as the DOS version had several solutions and names were refined re-designed and implemented in the new version. For the support of the novice users the On-line help facility is also programmed as to the Borland C++ language supports the Windows help system. The content of the help file should meet the specification of the Microsoft Help Compiler format. Missing or modified help screens will be added in English and/or translated to Hungarian by HCSO later on. Help screens may contain keywords, highlighted words, or so called hot spot item, on which the user can have more detailed information.

One of the most significant change in the system is the handling and the management of the area variable. It is now integrated to the other statistical variables. The user can create so called artificial variables using the selection function or grouping feature of the natural elements of the natural statistical indicators, and then and if is desired, one or more artificial variable can be set up from natural or grouped items. In this way various selections can defined and all user-defined artificial

variable can be used for layout and criteria definition.

The system can handle hierarchical categories, (such as the physical and intellectual in occupation, and within the physical occupations it should identify smoother sub-categories, for example in the industry, in agriculture, in traffic, in trade, etc., and in the industry are the mining, heavy industry, soft industry, etc., up to four levels.) The system can create and display multidimensional layout of cross-tabulated tables using the model layout, finding the variables in column and or the layer sections.

The variable definition and table composition is a very straightforward procedure and the user can always control the definition and then look the result. The user can have an instant respond about the development of the statistical table. Then, the retrieval offers the same function for the user, as it was earlier: from user defined areas by user defined structure of tables in one layout now, and by user defined criteria.

The other changes is the elimination of the spreadsheet module from the Windows version. Almost all users notes the unnecessary of the functions of the built-in spreadsheet program. It was reported that the commercial packages such as Lotus for Windows, Quattro Pro for Windows and Excel can manage the row-column oriented data. Once the data are extracted from the database on CD-ROM the aggregated elements can be transformed via the Lotus format to the packages referred above. If other would prefer dBASE or other XBASE packages and solutions for data processing they can transform the tables into dBASE formats. The other users who needs to process the data with a word processor the conversion to text format is supported. The fourth group of the users may want to process the data using statistical packages, so the one of the them, the SPSS format is supported.

The mapping part of the system was also redesigned: many customers do not use and did not feel the basic functionality of the creation of the basemaps. Many of the users requested a direct and fast link the numeric tables to map database. The design and the set up of the base map is now missing but extended formats are supported the management of graphics : Windows Meta File (WMF), bitmap (BMP) and DIB. In addition, the other business graphics, such as bar, pie and line can also be created directly from the tabulated data.

### Performance

The retrieval from the data stream has been speed up for the access of databases for multi-table queries and criteria. In many cases the speed of the retrieval is four or three times faster than the DOS version of SDBQ. Using the memory management of the Windows it is possible to create multi-columns tables which was one of bottlenecks of the previous version. The mapping system built in the software also uses the extended memory management and can support all country maps including all towns and villages. The DOS version had the limitation of mapping 1,000 geographical unit, so the entire country could not be presented on the same screen.

Using the Windows memory managements the prescribed variable limits for 256 is no longer valid, even more or up to 32,000 base statistical variables can be used; which is much more than anticipated or would be used in the future.

Printing of graphics inside the system is longer limited to a supported list of printers, in terms of special driver set. All printers are supported by Windows and once it is installed under Windows it should work properly.

## Programs and Coding

All programs were coded in Borland C++ version 4.5. The IDE project file enclosed to the programs contains the nodes, ie. module names and specification for compilation, and linking or binding the nodes for the target. The target file is EXE and now it is generated for Windows 3.1. Resource files (.RC) holds the specification of the windows, boxes, buttons and of all controls supported by Windows and the Borland. All windows, screens, messages are stored in different files according to different languages. These resources can be compiled and added to the program as Dynamic Link Libraries (DLLs).

Programs were coded in C++ using Object Oriented Programming (OOP) approach which is the latest development in programming technique against the conventional programming in the 1990's. Inheritance of the objects can be drawn using proper tools in the integrated development environment only. The interface and control programs of the application utilize the Borland Object Windows Library (OWL) version 2.0 provided with the Borland C++ compiler.

Header files contain the compiler directives, data and class structures and the program files (named with .CPP extension) has the class definitions and the application of the instances of the objects. Only one function is coded in assembler to get the highest performance. This is for fetch the records to be processed and it decompresses the given fields to numeric values. The critical function of the tabulation, such as building multi-dimensional array and functions of the classes are also optimized: it may use register and block variables, and use effective pointer arithmetic for the fast access of the data.