ACG SUB-COMMITTEE ON STATISTICAL ACTIVITIES

Twentieth Session
Geneva, 9 - 13 June 1980
Item 9 of the provisional agenda

ISSUES SURROUNDING THE USE OF MICRO-COMPUTERS
FOR STATISTICAL PURPOSES

Report prepared by the FAO
The use of microcomputers in statistical work in the Statistical Division/FAO.

At FAO, microcomputers are used both at the headquarters offices in Rome and in the field projects in various countries. FAO is a specialized agency and rarely originates the introduction of computer technology to the developing countries. Thus, computer facilities used for the statistical processing in the field, are limited to those already existing in the country. The types and makes of microcomputers used for data processing vary from country to country. Generally, the microcomputers lack statistical software packages.

In some field projects, only limited data processing tasks are conducted due to lack of suitable tabulation packages for that type of computer. One remedy for the lack of data processing packages in the field, is to enter the data locally. Then, the data is brought to headquarters and transferred to the mainframe for further tabulation using the SAS package.

SOFTWARE

Until recently, the use of microcomputers in the field was limited to data entry and simple tabulation. To cope with the limitations of locally available microcomputers, FAO's computer expert have mainly relied on ad hoc developed statistical software. In some countries, DRASS II and DATASTAR were used for the data entry and STATPAK for processing of the data. Preparation of the statistical data was done using spreadsheet packages such as SUPERCALC and LOTUS 1-2-3.

Until a few years ago, the statistical work with microcomputers was limited by the low capacity of microcomputer hardware. PC (microcomputer) versions of statistical mainframe packages such as SAS and SPSS require a microcomputer equipped with a hard disk. With the introduction of SPSS/PC and SAS/PC high quality general purpose statistical software is available for microcomputers.

The Statistics Division of FAO is trying to promote a certain degree of standardization of microcomputers for the statistical work in the field.

A large experimental project for the exchange of the aggregated statistical data in computer-readable form (i.e. diskette), was conducted by the Statistics Division. The experiment involved 11 countries, all having identical 8-bit microcomputers.

The data collection process is the largest source of error in the production of statistical data. In order to improve the quality of collected data, the pilot enumeration with the use of handheld microcomputers is being prepared.

The Statistics Division has identified several packages applicable for the traditional tasks of survey and census processing i.e. data entry, editing and tabulation.
The packages listed below are part of an integrated computing environment, which guarantees that the data can be transferred from one package into another.

**Processing chain 1**

```
<table>
<thead>
<tr>
<th>processing package</th>
</tr>
</thead>
<tbody>
<tr>
<td>steps</td>
</tr>
</tbody>
</table>

(INPUT) ------ -> (raw data)       
 ↓ (direction of data transfers)    
  ↓ 

data entry: ENTRYPOINTI
        or ROLL/PC
        ↓ 

위                         
correction: CONCOR
        ↓ 

tabulation: CENI 4
        ↓ 

(OUTPUT) <------ (tables)
```

**Processing chain 2**

```
<table>
<thead>
<tr>
<th>package</th>
</tr>
</thead>
</table>

(INPUT) ------ -> (raw data)       
 ↓ (direction of data transfers)    
  ↓ 

data entry: SAS (allows full screen design)
        or dBASE III
        ↓ 

위                         
correction: CONCOR
        or custom program
        or nothing (if data entry program in dBASE
        picks up all possible errors)
        ↓ 

tabulation: SAS/PC
        ↓ 

(OUTPUT) <------ (tables)
        ↓ 

further manipulation LOTUS 1-2-3
and presentation dBASE III
```

In the above chain SAS can be replaced by SPSS.
The SAS/PC can also be successfully used for the statistical...
HARDWARE

It is very time consuming and expensive to support incompatible hardware. Therefore in the future the Statistics Division would like to promote the use of hardware capable of running standard software packages.

At present, the largest spectrum of statistical software is available on IBM PC and compatible machines. It is this data processing environment, which the Statistical Division could easily support.

MICROCOMPUTER BASED STATISTICAL DATA PROCESSING ENVIRONMENT

The statistical data processing on PCs should, to the extent possible, satisfy the following criteria:

1. Use standard hardware.
2. Use software packages rather than rely on in-house programming.
3. PCs should be able to exchange data with the mainframe.
4. In the future, PCs should be integrated in the networks.
5. If possible, the same software should be used on PCs as on the mainframes.
6. The software packages on a single PC should provide an integrated environment so that the data can be easily moved between various packages.

PROBLEM AREAS

The lack of hardware standardization, maintenance and suitable software are the major problem areas dealt with by the backstopping staff. Especially, that the Statistics Division is frequently compelled to use already existing hardware. Frequently, software will not run on incompatible microcomputers and programming is rarely a feasible alternative.

By providing software and hardware guidance the Statistics Division expects that a certain degree of standardization can be achieved which will allow developing countries to benefit from the exchange of experience.

References: