CYSTAT: The road to modernisation

**Keywords:** GSBPM, Statistical Data Warehouse, Government Data Warehouse, Statistical Metainformation System, Statistical Information System

# Introduction

The overall organization of statistical activities in the Statistical Service of Cyprus (CYSTAT) is based on the traditional stove pipe model in which all the production processes are decentralized and carried out by the subject-matter Divisions/Sections. However, nowadays, there is an increasing number of National Statistical Institutes in which, the production processes are carried out in standardised procedures regardless of the content of the production output. As a result, there is increased efficiency and improved output quality.

A valuable tool in the specification of the standard production steps is the Generic Statistics Business Process Model (GSBPM) which CYSTAT started to implement in 2018 aiming to standardise the processes of the statistical production and abandoning the stove pipe model.

At the same time, the current technical infrastructure of CYSTAT faces several challenges. Thus, there is a need for the development of a statistical data warehouse to ensure the central storage and easy access to the statistical data and metadata.

CYSTAT is a small office with limited human and financial resources. Any reorganisation requires a significant amount of time to be invested in setting up the new structures and processes while at the same time the business continuity should not hampered. The development of a statistical data warehouse requires expertise which is not available inhouse. Thus, an opportunity for CYSTAT to modernise was brought up within the framework of the expansion of the Government Data Warehouse (GDW).

# Methods

In 2018, CYSTAT initialized two projects with the aim to modernise the statistical production processes. The first project deals with the use of the GSBPM to describe and define the set of business processes that are needed to produce official statistics. The second project concerns the development of a statistical data warehouse based on the technological infrastructure of the GDW. Both projects are under the umbrella of the Methodology and Quality Management Unit (MQMU).

## Implementation of the GSBPM in CYSTAT

The production units are asked to provide information for a selected number of 36 statistical surveys/ activities. The selection criteria applied ensure that a) there is information from almost all Sectors of CYSTAT, b) various types of surveys/ work are covered (e.g. business surveys, households, indicators etc) and c) none of the responsible officers is asked to fill out more than one questionnaire.

Due to the significant amount of time needed to provide descriptions for all sub-processes defined, the collection of information takes place in two stages. At a first stage, the production units are asked to provide information on selected sub-processes deemed to be of primary importance whereas the information on the remaining sub-processes are collected at the second stage.

The information is collected with the use of a questionnaire developed by the Methodology and Quality Management Sector of CYSTAT. The structure of the questionnaire is such that the official definition for each phase and sub-process is provided, followed by a set of questions aiming to cover all the relevant information.

## Development of a Statistical Data Warehouse

The Government Data Warehouse is a single cohesive database with a subject-centric approach, in order to provide a consolidated view of Civil Service data, optimized for reporting and analysis.

The Department of Information Technology Services (DITS) of the Cyprus Government decided the expansion of the GDW for the purpose of being used more extensively as the Singe Version of Truth (SVT) containing in this way consistent and accurate information about government’s departments business reality. Part of this expansion is the Statistical Metainformation System (SMS) of CYSTAT which aims to ensure a systematic utilization of Metainformation, both inside and outside of the Statistical Information System (SIS). The goal is to be able to integrate efficiently the SMS and ensure internal and external addition to the GDW expansion.

The requirements for the expansion of the GDW were defined by a team of experts. As regards the statistical data warehouse the experts identified the bottlenecks of the current stove pipe model, defined the aims and prepared the requirements. All the information was gathered through the available documentation, interviews with the personnel of CYSTAT and the close cooperation of the MQMU which also had a catalytic role in the analysis.

# Results

Both are long term projects, however, the preliminary analysis carried out so far identified the bottlenecks of the current statistical information system. As regards the statistical production, the analysis confirmed the need to move on to a new production model base on standardized procedures. This will inevitably have consequences in the organisational structure. A major strategic ail of CYSTAT is the increased use of administrative sources. In this respect, CYSTAT should ensure the continuous access to the various sources and thus, the integration of CYSTAT’s information system with the GDW is of primary importance.

Moreover, the analysis showed that diverse application software packages are used and that, there is a lack of central data storage. Subject matter statisticians have to spend a significant amount of time by performing technical work. The consequence is that, they are lacking time and time and capacity for the substantial, content oriented, statistical and methodological activities.

As regards metadata, the preliminary analysis showed that there is a lack of coordination of metadata description for CYSTAT’s statistical data. There is no sufficient coordination in the definitions on statistical indicators/ variables. Statistical classifications are maintained in the responsibility of individual-subject matter Divisions/Sections. There is no central administration of statistical classifications in CYSTAT. In the time being, a coherent statistical metainformation system does not exist in the CYSTAT. Its development is a key prerequisite for the innovation of SIS.

# Conclusions

The proliferation of statistics, produced by CYSTAT, calls for efficient methods and techniques for storage and maintenance of statistical data. Advantages of data warehouse technologies meet CYSTAT requirements. Uniformity of data storage and related metadata description, easy and quick access to data, powerful possibilities of data export towards users and a potential for data protection are examples of such important data warehouse features. The development of SMS will contribute to upgrade methodological work in CYSTAT, to increase efficiency of statistical activities and to enhance quality of statistical data/information.

A significant aspect of the modernisation project is the fact that the statistical data warehouse and SMS will be built based on the infrastructure of the GDW and in this way, any technological conflicts will be avoided.

The modernisation is an ambitious and long-term project. It is imperative that for the project to be successful the support of the top management and the commitment of the personnel are perhaps the most critical factors.