

Multiple Sources in Business Surveys

Paulo Saraiva, Almiro Moreira and Sofia Rodrigues (Statistics Portugal)

paulo.saraiva@ine.pt; almiro.moreira@ine.pt; sofia.rodrigues@ine.pt

Abstract

Intensive use of multiple data sources is a strategic action in the context of producing official statistics. However, it is not a trivial matter to deal with alternative data sources. In fact, the tradition of official statistical production persists largely on surveys, in which many factors of change management prevail.

Statistics Portugal has extensive experience in a centralized approach to data collection as well as an integrated survey management system. This system, based on components of statistical units and metadata, also adopts a data-centric approach, which favors the expansion of the use of administrative sources.

This paper presents Statistics Portugal's experience in business surveys using data from various sources, guided by the Generic Business Statistical Process Model (GSBPM), particularly in the Specify Needs, Design and Collection phases. Relevance is given to the "check data availability" sub process, where a tool such as a catalogue of existing or potential administrative data sources is critical and is being developed. Also discussed are the efforts required for process and data integration, where the tasks of combining data from multiple sources, record matching routines, and prioritization are essential, especially when two or more sources contain data for the same variable.

Keywords

administrative data; statistical data management; potential administrative data usage; data collection; business surveys; official statistics