

IMPUTATION METHODS OF ITEM NON-RESPONSE IN EU-SILC

Ruzhdie Bici, PhD¹

¹ *Senior Statistician, Sogeti, Luxembourg, ruzhdie.bici@sogeti.com¹*

Résumé. Les non-réponses et les informations manquantes sont inévitables dans les enquêtes. Ces données manquantes conduisent à des estimations biaisées et des résultats non fiables. Il est donc préférable pour ces raisons de recourir à l'imputation statistique pour les corriger.

Dans cet article, nous nous concentrerons sur la non-réponse partielle et les différentes méthodes utilisées pour y répondre. L'analyse est basée sur les résultats 2023 de l'enquête de EU-SILC et sur les pratiques nationales des pays réalisant l'enquête.

Les pays ont choisi d'appliquer différentes techniques, telles que l'imputation logique ou par donneur, mais aussi des modèles statistiques plus complexes pour imputer les données manquantes. La source d'information, le mode de collecte des données, le type de variable et les informations disponibles pour cet enregistrement sont considérés comme des facteurs pouvant influencer les méthodes utilisées par les pays. Ces facteurs et les différentes méthodes appliquées par les pays afin de réduire les informations peu fiables et les résultats biaisés sont ainsi analysés.

L'article fournit des recommandations utiles sur la manière de traiter la non-réponse partielle dans EU-SILC ainsi que pour d'autres enquêtes par sondage lors de leur future mise en œuvre.

Mots-clés. Non-réponse partielle, enquêtes, imputation, EU-SILC, données manquantes

Abstract. Non-response and missing information are unavoidable in surveys. Missing data leads to biased estimations and unreliable results. For this reason, it is preferable to use statistical imputation.

In this paper, the focus will be on item non-response and different methods used to decrease item non-response. The analysis is based on the national practices and results of the EU-SILC 2023 data.

Countries chose to apply different techniques, like logical or donor imputation, and as well as more complex statistical models to impute missing data. The source of information, the data collection mode, the variable type, and the information available for that record are considered as possible factors impacting the method used by countries. These factors and different methods applied by countries to decrease unreliable information and biased results are analysed.

The paper provides useful recommendations on dealing with item non-response in the EU-SILC and in other sample surveys in future implementation.

Keywords. Item non-response, surveys, imputation, EU-SILC, missing data