



CONFERENCE OF EUROPEAN STATISTICIANS

WP.16

Workshop on Statistical Data Collection: Riding the Data Deluge

23 April 2015

29 April – 1 May, Washington D.C., United States of America

United Nations Statistics Division (UNSD)'s Big Data Pilot Project Initiatives

Nancy Snyder (United Nations Statistics Division)

snydern@un.org

Abstract

At its 45th session last year, the Statistical Commission recognized the importance of Big Data for official statistics and established a Global Working Group to develop guidance in this area. Big Data are potentially useful for official statistics, but require new tools and methods for capturing, managing and processing them efficiently. To take full advantage of these data sources, the statistical community needs to adequately address issues pertaining to methodology, quality, technology, data access, legislation, privacy, management and finance, and provide adequate cost-benefit analyses. The United Nations Statistics Division (UNSD) encourages compilers to undertake pilot projects in using Big Data to explore their potential use in compiling official statistics, in combination with traditional data sources, in order to improve timeliness, fill data gaps, and reduce cost, while maintaining a high-level of quality resulting official statistics. UNSD is partnering with certain countries and organizations to initiate pilot projects in the use of Big Data for official statistics. The national statistical office of Costa Rica and the UNSD linked Costa Rica's existing Customs-based trade statistics and its business register, which is maintained from administrative data sources, specific projects and follow-up telephone interviews. As this experience shows, new micro-level data, new statistical indicators and new insights (for instance about the international orientation of businesses and about trade by enterprise characteristics) can be gained on a continuing and routine basis through this linking procedure, which can be applied in its methods and outcomes in other developing countries. The next phase of this project will be further linking Costa Rica's trade and business data with postal data, both at the national level and international level, from the Universal Postal Union (UPU). Such linking will allow for exploring the use of Big Data from UPU as a proxy or leading indicator for trade volumes; for data mining the microdata for information on the volume of international trade at the local level, the speed of international trade, and transportation issues; and for possibly estimating selected targets for the sustainable development goals (SDGs). UNSD is also working with Oman to better understand the nature and potential statistical applications of mobile phone data, including statistics on tourism (both domestic and international), migration, and population, among others. UNSD is also collaborating with national statistical offices and other organizations on the UNECE Big Data Sandbox in order to explore tools and methods needed for statistical production and the feasibility of producing Big Data-derived statistics. On this platform, UNSD is contributing highly-detailed and disaggregated international merchandise trade statistics from its global database, UN Comtrade, in order to provide opportunities for new research, including regional value chain analysis, analysis of bilateral trade asymmetries, and calculation of trade indices, among others.

