Country Statement – Malaysia

ITEM 4(b): EMBRACING BIG DATA TO MODERNIZE OFFICIAL STATISTICS IN ASIA AND THE PACIFIC

Regional and national efforts to embrace big data for official statistics in Asia and the Pacific

Malaysia would like to share its experience in big data development. The Department of Statistics Malaysia (DOSM) has embarked on Big Data Project since December 2016. As a start, three (3) non-traditional data sources were explored which include online price data, media social data and online news data. An adoption of web crawling techniques price data collection tools in order to identify the potential use in statistical production process (i.e. the calculation of Consumer Price Index). Currently, online price data is used internally to monitor price movement and historical price trends. During COVID-19 pandemic, price web crawling method has become very useful since the enumerators cannot reach the outlet to collect price data. The data that obtains from online social media i.e. Twitter & Facebook is used to analyse and assess public opinion with regards to official statistics published by DOSM. On top of that, online news data also can be accessed through the same platform. Latest data news and trends, top person mentioned in the news, top influencer on particular topics and etc. can be analysed by providing related keywords. Besides that, in early 2020 a new collaboration between DOSM with remote sensing agency has been set up to explore on satellite imagery data for land used statistics and agriculture statistics. It is also used to explore the potential of satellite data for rural area during 2020 Population Census. This satellite imagery data will be studied to identify how it can be used to improve official statistics on a wide range of topics encompassing agriculture, the environment, and socioeconomic activity.

Therefore, National Statistical Office (NSO) is encouraged to form a new international partnership to leverage on the various sources of data for producing official statistics. It could be as simple as working together on a cleansing algorithm so partners can benefit from clean data. Private sectors can spend on cost, open up private data and share IT skills.