Economic and Social Commission for Western Asia

Modernizing Collection Methods for Price Statistics in the Arab Region

UNITED NATIONS
الاسكوا
ESCWA

Shared Prosperity **Dignified Life**

Majed Skaini

TAGES MEETING
6-7 July 2021

Introduction

ESCWA is always looking for new opportunities to develop and modernize price statistics.

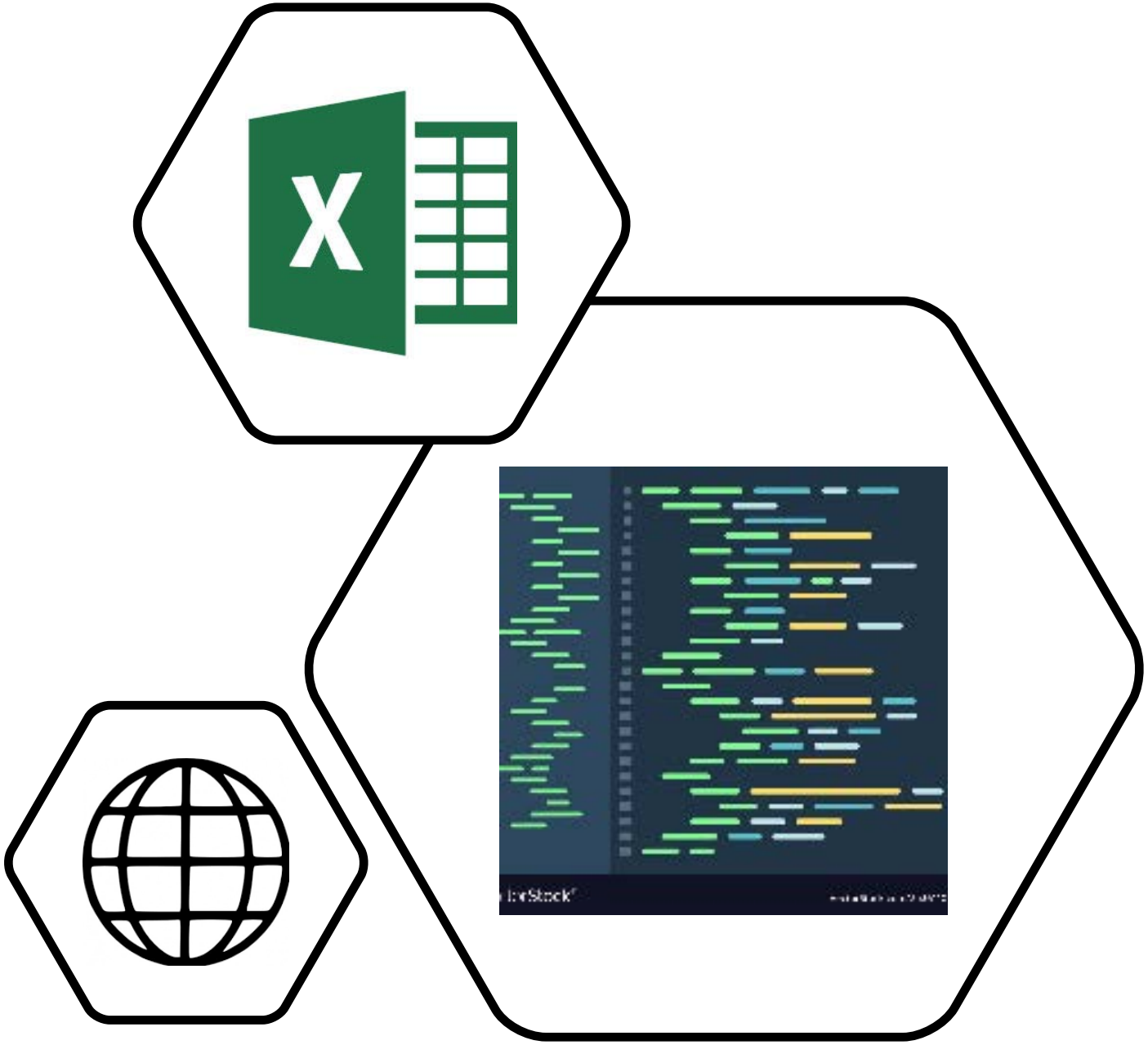
With the advance in technology and the prevalence of Big Data, ESCWA started looking for ways to apply Big Data tools in price statistics to improve statistical output and quality.

Big Data tools explored by ESCWA for use in price statistics include Web Scraping and Scanner Data as means of complementary, non-traditional, data collection methods.

I. Web Scraping

What is Web Scraping?

- Web scraping is a data collection method that automatically extracts large amounts of data from websites and saves data to a local file or database in a structured format.



Web Scraping Efforts in the Arab Region

ESCWA started researching web scraping and its application in price statistics in 2019.

ESCWA started piloting web scraping in a few countries in early 2020, namely Bahrain, Kuwait, Lebanon and Qatar.

The pilot application focused on scraping names and prices for technological items - such as laptops and smart phones - from a few large outlets.

After the success of the pilot phase and the need for non-traditional, centralized, data collection with the spread of COVID-19, ESCWA widened the scope of its web scraping application.

Web Scraping Efforts in the Arab Region (cont'd)

ESCWA developed new scrapers for Bahrain for different household consumption categories with requested additions.

New improvements to scrapers include product description, SKU code and product availability when such details are available in outlet websites.

Training workshops were conducted for Jordan and Egypt during June 2021, transferring the know-how and preparing national statistical office staff to write scrapers themselves.

Scrapers have also been developed for several additional countries, namely Saudi Arabia, Oman, Tunisia and the UAE and training sessions will be conducted for these countries.

How does ESCWA develop web scrapers?

1

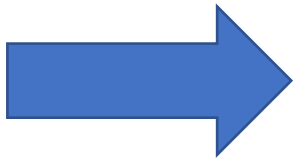
Member countries provide ESCWA with the links of reliable and big outlet websites to collect data from.

2

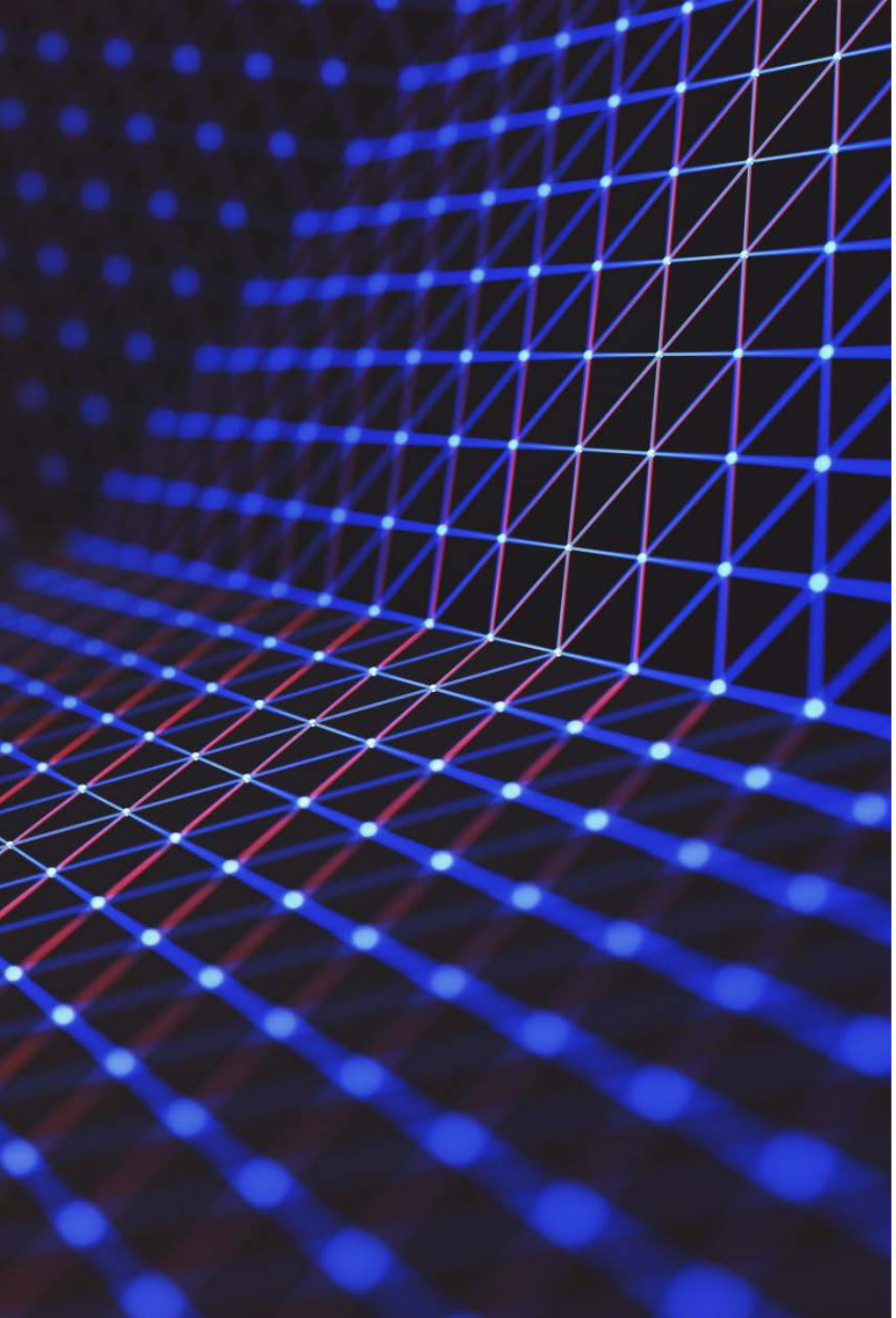
ESCWA develops its scrapers from scratch, writing the software using Python programming language.

3

Requested data is automatically extracted from the outlet website to an excel file with one click of a button.



Currently, ESCWA is looking into the incorporation of additional features, such as automation and methods for linking outlet codes to COICOP codes to facilitate the use of scraped data.



Benefits of Web Scraping over traditional data collection

- Automation
- Centralization
- Higher frequency of data
- Bigger volume of data
- Higher efficiency
- Higher accuracy

However, web scraping does not completely replace traditional data collection and field visits but provides a <u>complementary</u> data source.

Country Feedback

This initiative received a lot of positive feedback from the participating countries which have already started utilizing and benefiting from the developed scrapers in the collection of price data for the computation of the CPI and other price indices.

The participating countries are also requesting additional training and workshops to take this initiative another step further.

II. Scanner Data

What is Scanner Data?

Scanner data is detailed data on sales of consumer goods obtained by scanning the barcodes for individual products at electronic points of sale in retail outlets, such as supermarkets for example.

It can provide detailed information about quantities, characteristics and values of goods sold as well as their prices for a given period.

ESCWA's Efforts in the Application Scanner Data

ESCWA has been researching scanner data and its applicability in the Arab region.

The following roadmap was put in place for utilizing scanner data:

- Choosing the largest retailers with electronic points of sale.
- Establishing agreements and legal framework.
- Reaching an agreement with outlets on data transmission mechanism and data requirements.
- Setup of the data transmission software.
- Technical training of staff.

Benefits of Scanner Data

- Scanner data provides multiple benefits to national statistical offices.
- Other than the benefits provided by web scraping, additional benefits include the availability of quantities and expenditures at product levels and the availability of back series of data.

Challenges in the Application of Scanner Data

Outlet reluctance to provide the requested data

Delays in data transmission

Issues in the scope of data coverage

Data processing difficulties and need for technical IT skills

Cost of setting up data transfer software

Scanner Data in the Arab Region

Some Arab countries, such as Oman, have been utilizing scanner data for a while.

Other countries, namely Egypt, have started the preparation process and are currently working on the legal aspect and legislations.

ESCWA provides guidance in the steps needed to apply scanner data, however legal and other agreements between NSOs and outlets have to be made at the national level.

UNITED NATIONS
الاسكوا
ESCWA

Shared Prosperity Dignified Life

Thank you