# U.S. BUREAU OF LABOR STATISTICS

## THE CONSUMER PRICE INDEX

# Modernizing Data Collection: APIs & Web Scraping

To continue to produce high quality data, the CPI is looking to supplement traditional survey collection with more modernized methods. While the BLS already includes prices for goods and services purchased online, the CPI is exploring APIs and web scraping - extracting data directly from websites. APIs and web scraping are efficient and cost-effective methods of collection that will make it even easier for businesses and organizations to participate in CPI surveys.



## What are APIs?

Some establishments provide Application Programming Interfaces (APIs) to allow partners to access information on their website. Data collection through an API is often easier and more straight-forward than maintaining web scraping code over time. By entering certain parameters, such as SKU code and store address, APIs can identify specific products and prices.



# What is web scraping?

Web scraping is a process through which information is gathered and copied from the web for analysis. Web scraping uses software that simulates human web surfing to collect existing information from a company's website without website disruption.



# Why use APIs and web scraping?

- Enhance the ease of participating in CPI surveys by saving businesses and organizations time and resources.
- Utilize high-speed methods to acquire large volumes of information amassing more data than staff collecting

## What is the CPI?

The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by urban consumers for a market basket of goods and services. The CPI is one of the most closely watched Principal Federal Economic Indicators (PFEIs) produced by the U.S. Bureau of Labor Statistics (BLS). Some of the people who follow the CPI closely include landlords (to help them calculate changes in rental prices), Social Security recipients (to anticipate their next cost-of-living adjustment), and Wall Street financial institutions and traders (to determine potential moves of the stock and bond markets).

## Your Voluntary Cooperation

Data collection for the CPI will not pursue any web scraping activities unless permission has been granted. The agency will engage in this activity in a responsible and transparent manner. BLS will take steps to make sure CPI data collection programs have minor impact on the website by limiting the time of day to execute web scraping programs, number of fetches per hour/day, time interval between requests, and type of data extracted. This report is authorized by law, 29 U.S.C.2. Your voluntary cooperation is needed to make the results of this survey comprehensive, accurate, and timely.

# Your Confidential Participation

The Bureau of Labor Statistics, its employees, agents, and partner statistical agencies will use the information you provide for statistical purposes only and will hold the information in confidence to the full extent permitted by law. In accordance with the Confidential Information Protection and Statistical Efficiency Act (44 U.S.C. 3572) and other applicable Federal laws, your responses will not be disclosed in identifiable form without your informed consent. Per the Federal Cybersecurity Enhancement Act of 2015, Federal information systems are protected from malicious activities through cybersecurity screening of transmitted data.

## How can I see CPI data?

Information is always available on the BLS-CPI homepage at www.bls.gov/cpi/home.htm.

New CPI data appear in a news release usually issued between the 10th and 15th of the month, reporting the data for the previous month. Also at this time, CPI data is reported in various media, such as television, newspapers, and public websites.



# U.S. BUREAU OF LABOR STATISTICS

# THE CONSUMER PRICE INDEX

# BLS national and regional offices

### Washington, DC

Bureau of Labor Statistics 2 Massachusetts Avenue, NE Washington, DC 20212 (202) 691-7000 CPI\_info@bls.gov

#### **Atlanta**

Bureau of Labor Statistics 61 Forsyth Street, SW, Room 7T50 Atlanta, GA 30303 (404) 893-4222 BLSInfoAtlanta@bls.gov

#### **Boston**

Bureau of Labor Statistics JFK Federal Building, E-310 Boston, MA 02203 (617) 565-2327 BLSInfoBoston@bls.gov

### Chicago

Bureau of Labor Statistics J.C. Kluczynski Federal Office Building 230 South Dearborn Street, Room 960 Chicago, IL 60604 (312) 353-1880 BLSInfoChicago@bls.gov

#### **Dallas**

Bureau of Labor Statistics 525 South Griffin Street, Room 221 Dallas, TX 75202 (972) 850-4800 BLSInfoDallas@bls.gov

#### **Kansas City**

Bureau of Labor Statistics Two Pershing Square Building 2300 Main Street, Suite 1190 Kansas City, MO 64108 (816) 285-7000 BLSInfoKansasCity@bls.gov

#### **New York**

Bureau of Labor Statistics New York-New Jersey Information Office 201 Varick Street, Room 808 New York, NY 10014 (646) 264-3600 BLSInfoNY@bls.gov

## Philadelphia

Bureau of Labor Statistics Suite 610 East – The Curtis Center 170 South Independence Mall West Philadelphia, PA 19106 (215) 597-3282 BLSInfoPhiladelphia@bls.gov

#### San Francisco

Bureau of Labor Statistics 90 7th Street, Suite 14-100 San Francisco, CA 94103 (415) 625-2270 BLSInfoSF@bls.gov

## **Questions?**

If you have any questions or comments regarding any aspect of the survey, you may contact:

Division of Consumer Prices and Price Indexes

Bureau of Labor Statistics 2 Massachusetts Avenue, NE Washington, DC 20212 call (202) 691-6991, or email: **CPI info@bls.gov**.

The U.S. Office of Management and Budget (OMB) has approved this collection of information and has assigned 1220-0039 as the control number. Without OMB approval and this number, we would not be able to conduct this survey.

