

Consumer Expenditure Surveys LABSTAT Database Getting Started Guide

Aaron Cobet

April 4, 2019

Consumer Expenditure Surveys

U.S. Bureau of Labor Statistics

Introduction

This document provides an overview of the CE LABSTAT database. The data are available through the following search tools:

- **Data Finder** allows you to search for selected time series variables by characteristics, create charts and tables, and export the data into Excel and CSV.
- **One-Screen Data Search** allows you to produce customized time series by households and their characteristics. Data can be exported into Excel through a single-screen form.
- **Multi-Screen Data Search** allows you to produce customized time series by households and their characteristics. Data can be exported into Excel through a series of steps using a multi-screen form.

To use these tools, select one and click through the options. The search tool provides the selected data and the corresponding series ID. You can generate time series by entering the Series ID into the [Series Report](#).

Not all tabular data are in the CE LABSTAT database. For example, The CE LABSTAT database only contains annual data, so Midyear tables cannot be replicated using the database. Similarly, the CE LABSTAT database does not include data on variances or by metropolitan statistical areas. You can find the additional tabulations on the [CE tables page](#).

Table of Contents

Section 1: Survey description and data overview	1
Section 2: Description of codes, series IDs, data, and administrative files	2
Section 3: Series file format and field definitions	3
Section 4: Data file format and field definitions	4
Section 5: Mapping file formats and field definitions	4
Section 6. Data element dictionary	6

Section 1: Survey description and data overview

Survey description: The CE surveys collect information from the Nation's households and families on their buying habits, income, and other demographic characteristics. For more information, see the Handbook of Methods [Consumer Expenditures and Income page](#).

Data overview: This document provides an overview of the codes for the One-screen Data Search and the Multi-screen Data Search, which are useful applications for creating multi-year data sets and graphs for specified variables. Instead of copying and pasting values one year at a time from the CE tables into an Excel spreadsheet to cover a number of years, the databases allow users to retrieve all available years at once. Using these databases, one can create Excel data sets for selected variables. Users can specify the range of years for the data then graph the information for a trend analysis.

Summary data available: Data are sorted into four categories: Expenditures, income and taxes, consumer characteristics, and addenda containing assets and liabilities, and other financial information.

Average annual expenditures consist of 14 main categories: Food, alcoholic beverages, housing, apparel and services, transportation, health care, entertainment, personal care, reading, education, tobacco products and smoking supplies, miscellaneous, cash contributions, and personal insurance and pensions.

The average annual expenditures, income, characteristics per consumer unit, and addenda are classified by 14 characteristics, which correspond to 14 of the published CE standard tables: Income before taxes, quintiles of income before taxes, deciles of income before taxes, age of reference person, size of consumer unit, composition of consumer unit, number of earners, housing tenure and type of area, race of reference person, Hispanic or Latino origin of reference person, region of residence, occupation of reference person, education of reference person (available through 2012), and highest education level of any member (beginning in 2012).

Technical notes:

- Periodicity: The CE databases provide calendar year averages. In addition, the [CE tables page](#) provides mid-year tables with data covering July of one year to June of the next year.
- Base periods: CE data do not have base periods.
- Updating schedule: The databases are updated annually, usually in September.
- Data characteristics: Average annual expenditures and income are reported in dollars and are stored with no decimal places. Percent distributions are stored with no decimal places. Number of Consumer Units is reported in thousands and is stored with no decimal places. Consumer Unit characteristics are reported averages and are stored with one decimal place.

For more information, see the [BLS Handbook of Methods](#), and for questions, [contact CE](#).

Section 2: Description of codes, series IDs, data, and administrative files

This section lists files that describe codes to generate time series of most CE information found on the [CE tables page](#). In addition, the section lists files that contain series IDs, data, and general information related to the database.

- **cx.category** lists the codes for CE's main data types: Expenditures, income and taxes, assets and liabilities, and consumer characteristics.
- **cx.subcategory** lists codes for subcategories, including apparel, food, health, housing, transportation, and income.
- **cx.item** lists item codes. For example, the CE program provides apparel items, such as women and girls, or apparel for girls 2 to 15.

- **cx.series** file contains a set of codes which, together, comprise a series identification code that serves to uniquely identify a single time series, such as "Women, 16 and over by Income Deciles: Ninth 10 percent."
- **cx.data.1.AllData** is a data file that lists Series IDs, years, periods, and data values.
- **cx.demographics** lists codes for major demographic breakouts, such as income by quintile, or age of reference person.
- **cx.characteristics** lists the codes for each demographic category. For example, for income by quintile, the file lists all five quintiles separately and all consumer units together.
- **cx.footnote** lists the footnotes of the published CE tables.
- **cx.process** provides the code to identify the data provided by this database as annual means for all households.
- **cx.contacts** provides CE program contact information.

Section 3: Series file format and field definitions

The following file format is used to define series files. Series files are in American Standard Code for Information Interchange (ASCII) text format, which is a common format for text files. Data elements are separated by tabs and the first record of each file contains the column headers for the data elements stored in each field. Each record ends with a new line character.

Below are the types of fields, their length, and example values.

Data element	Length	Example value
series_id	30	CXUFRSHFRUTLB0201M
seasonal	1	U
category_code	10	EXPEND
subcategory_code	9	FOOD
item_code	10	FRSHFRUT
demographics_code	4	LB02
characteristics_code	2	01
process_code	1	M
series_title	256	It varies
footnote_code	10	It varies
begin_year	4	It varies
begin_period	3	AO1
end_year	4	It varies
end_period	3	AO1

The series_id CXUFRSHFRUTLB0201M can be broken out into these segments:

Data element	Example value
survey abbreviation	CX
item_code	FRSHFRUT
demographics_code	LB02
characteristics_code	01
process_code	M

Section 4: Data file format and field definitions

The following file format is used to define data files. Data files are in ASCII text format. Data elements are separated by tabs. The first record of each file contains the column headers for the data elements stored in each field. Each record ends with a new line character.

Below are the possible types of fields, their length, and values for the data file "cx.data.1.AllData."

File name: cx.data.1.AllData

Data element	Length	Example value
series_id	30	CXUMENBOYSLB0104M
year	4	1984
period	3	A01
value	12	93
footnote_code	10	It varies

The series_id CXUMENBOYSLB0104M can be broken out into these segments:

Data element	Example value
survey abbreviation	CX
seasonal	U
item_code	MENBOYS
demographics_code	LB01
characteristics_code	01
process_code	M

Section 5: Mapping file formats and field definitions

The following file format is used to define the mapping files. Mapping files are in ASCII text format. Data elements are separated by tabs. The first record of each file contains the column headers for the data elements stored in each field. Each record ends with a new line character.

File name: cx.category

Data element	Length	Example value
category_code	10	EXPEND
category_text	60	Expenditures
display_level	2	0
selectable	1	T
sort_sequence	5	100

File name: cx.subcategory

Data element	Length	Example value
category_code	10	EXPEND
subcategory_code	9	PERSCARE
subcategory_text	50	Personal care products and services
display_level	2	0
selectable	1	T
sort_sequence	5	900

File name: cx.item

Data element	Length	Example value
subcategory_code	9	ENTRTAIN
item_code	10	PETSPLAY
item_text	50	Pets, toys, and playground equipment
display_level	2	1
selectable	1	T
sort_sequence	5	11200

File name: cx.characteristics

Data element	Length	Example value
demographics_code	4	LB01
characteristics_code	2	01
characteristics_text	120	All consumer units
display_level	2	1
selectable	1	T
sort_sequence	5	1010

File name: cx.demographics

Data element	Length	Example value
demographics_code	4	LB01
demographics_text	150	Quintiles before taxes
display_level	2	0
selectable	1	T
sort_sequence	5	100

File name: cx.footnote

Data Element	Length	Example value
footnote_code	2	It varies
footnote_text	250	No data provided

File name: cx.process

Data element	Length	Example value
process_code	1	M
process_text	30	Means
display_level	2	0
selectable	1	T
sort_sequence	5	100

File name: cx.seasonal

Data element	Length	Example value
seasonal	1	U
seasonal_text	30	Seasonally unadjusted

File name: cx.year

Data element	Length	Example value
year	4	2001

Section 6. Data element dictionary

This section describes the data elements in the CE LABSTAT database.

Data element	Length	Value example	Description
begin_period	3	A01 (A01 = Annual data)	Identifies first data observation within the first year for which data are available for a given time series.
begin_year	4	1999	Identifies first year for which data are available for a given time series.
category_code	10	EXPEND	Identifies the category code for the data series.
category_text	60	Text	Contains the category text.
characteristics_code	2	01	Identifies the characteristics code for the data series.
characteristics_text	120	Text	Contains the text of the characteristics.
display_level	2	0 (0 = No indentation)	Displays hierarchical data structure with indentations to the area descriptions.
end_period	3	A01 (A01 = Annual data)	Identifies the last data observation within the last year for which data are available for a given time series.
end_year	4	2002	Identifies the last year for which data are available for a given time series.
footnote_code	2	It varies	Identifies footnote for the data series.
footnote_text	250	Text	Contains the text of the footnote.
item_code	10	PETSPLAY	Identifies item for which data observations pertain.
item_text	50	Text	Describes item.
period	3	A01 (A01 = Annual data)	Identifies period for which data are observed.
process_code	1	M (M = Means)	Identifies process for which data are observed.
process_text	30	Text	Contains the text of the process.
seasonal	1	U (U = Unadjusted)	Identifies whether the data are seasonally adjusted.
selectable	1	T (T = Can be selected)	Distinguishes standard data from title information.
series_id	30	CXU080110LB0101M	Identifies specific series.
sort_sequence	5	100	Identifies the order in which the items should be displayed.
subcategory_code	9	APPAREL	Identifies the subcategory code for the data series.
value	12	275	Lists the observation for the series.

Data element	Length	Value example	Description
year	4	2001	Identifies year of observation.