

Consumer Expenditure Surveys Public-Use Microdata (PUMD) Overview

Jimmy Choi

Division of Consumer Expenditure Surveys
CE Microdata Users' Workshop
July 17-19, 2019

What Are Microdata?

- “In the study of survey and census data, microdata are information at the level of individual respondents.” – Wikipedia
- Consumer Expenditure Surveys (CE) Public-Use Microdata (PUMD) consists of Consumer Unit (CU) level characteristics, income, taxes, assets and liabilities, and expenditure data, as well as member level characteristics and income data.
- Data are provided through various file types.

Consumer Expenditure Surveys



BROWSE CE

- CE HOME
- CE OVERVIEW
- CE NEWS RELEASES
- CE PUBLICATIONS
- CE TABLES
- CE LABSTAT DATABASE
- CE PUBLIC USE MICRODATA
- CE WORKSHOP AND SYMPOSIUM
- CE GEOGRAPHIC DATA
- CE EXPERIMENTAL RESEARCH PRODUCTS
- CONTACT CE

SEARCH CE



CE TOPICS

- INFORMATION FOR CE RESPONDENTS
- CE LIBRARY
- CE METHODOLOGY
- CE SURVEY MATERIALS
- CE FAQ
- CE GLOSSARY
- CE REDESIGN
- CE DATA QUALITY AND

Public-Use Microdata (PUMD)

The Consumer Expenditure Survey (CE) collects data on expenditures, income, and demographics in the United States. The public-use microdata (PUMD) files provide this information for individual respondents, without any information that could identify respondents. PUMD files include adjustments for information that is missing because respondents were unwilling or unable to provide it. The files also have been adjusted to reduce the likelihood of identifying respondents, either directly or through inference.

This format allows researchers to analyze expenditures, income, and demographic trends beyond what the published tables show. However, users of the PUMD files should have strong skills with statistical software. To learn more, you may want to explore the PUMD methodology with the [CE PUMD Getting Started Guide](#) and check the availability of data with the [Dictionary for Interview and Diary Surveys \(XLSX\)](#).

The 2017 PUMD were released on September 11, 2018.

Note: To users of the CE 2017 Public-use Microdata: some corrected files have been reissued on June 27, 2019. See [Errata notice](#) for details.

[Experimental State Weights for 2017](#), supporting state level analysis for selected states, are also available with their [documentation](#).

To be notified when new datasets become available, please sign up for the [CEX e-mail updates](#). If you have comments, suggestions, and questions about this page and its contents, please [contact us](#).

Quick Links

- » [Introduction to PUMD](#)
- » [Data files](#)
- » [Documentation](#)
- » [Supplemental resources](#)

- » [Public-Use Microdata Annual Users' Workshop](#) provides users the opportunity to learn how to use the PUMD and to present how they used PUMD in their research.

<https://www.bls.gov/cex/pumd.htm>



PUMD Data Package

■ Data files:

- ▶ File types available include:

SAS (*.sas7bdat), STATA (*.dta), SPSS (*.sav), comma-delimited (*.csv)

■ Documentation:

- ▶ Documents covering specific years (changes with each new PUMD release)

- ▶ Documents covering all years (Static or evolving documentation)

■ Supplementary resources:

- ▶ Provide information regarding PUMD disclosure methods and requirements

Consumer Expenditure Surveys



BROWSE CE

- CE HOME
- CE OVERVIEW
- CE NEWS RELEASES
- CE PUBLICATIONS
- CE TABLES
- CE LABSTAT DATABASE
- CE PUBLIC USE MICRODATA
- CE WORKSHOP AND SYMPOSIUM
- CE GEOGRAPHIC DATA
- CE EXPERIMENTAL RESEARCH PRODUCTS
- CONTACT CE

SEARCH CE

 [Go](#)

CE TOPICS

- INFORMATION FOR CE RESPONDENTS
- CE LIBRARY
- CE METHODOLOGY
- CE SURVEY MATERIALS
- CE FAQ
- CE GLOSSARY

PUMD Data Files

From 1996 forward, public-use microdata (PUMD) are available in four formats: SAS, SPSS, STATA and Comma Delimited (ASCII). Each format presents the data for each year in two zipped files. Data collected with the Interview Survey are in the "Interview" file and those with the Diary Survey are in the "Diary" file.

Pre-1996 PUMD are available as Comma Delimited files. The [Inter-university Consortium for Political and Social Research](#) (ICPSR) provides the data for free to its members. Here is their [member list](#). BLS also sells the data on [USB flash drives \(PDF\)](#).

If you are new to the CE PUMD, you may want to explore the PUMD methodology with the [CE PUMD Getting Started Guide](#) and the availability of data with the [Dictionary for Interview and Diary Surveys \(XLSX\)](#). For those using income data in their analyses, we also recommend reading the [User's Guide to Income Imputation in the CE](#). Additionally, we supply training programs and sample data ([ZIP](#)), and host a free [microdata data users' workshop](#).

[Experimental State Weights](#), supporting state-level analysis for selected states, are also available with their [documentation](#).

The 2017 PUMD was released on September 11, 2018. To be notified when new datasets become available, please sign up for the [CEX e-mail updates](#). If you have comments, suggestions, and questions about this page and its contents, please [contact us](#).

If you are having difficulty downloading the data, please see the [instructions on how to download zip files](#).

SAS | SPSS | STATA | CSV

SAS

2017 | [Interview \(zip\)](#) (89.1 MB) | [Diary \(zip\)](#) (10.6 MB) | [Documentation](#)

2016 | [Interview \(zip\)](#) (91.6 MB) | [Diary \(zip\)](#) (10.5 MB) | [Documentation](#)

2015 | [Interview \(zip\)](#) (87.7 MB) | [Diary \(zip\)](#) (10.5 MB) | [Documentation](#)

https://www.bls.gov/cex/pumd_data.htm



Data Files

- PUMD files contain respondent data collected in either the *Interview* or *Diary* survey.
- These data can be used to create and correlate information in a variety of ways:
 - ▶ Descriptive Statistics
 - How many CU's in the sample purchased a new car in 2014?
 - ▶ Inferential Statistics
 - What is the average annual expenditure on eggs by all CU's in the northeast region?



Consumer Expenditure Surveys



BROWSE CE

[CE HOME](#)

[CE OVERVIEW](#)

[CE NEWS RELEASES](#)

[CE PUBLICATIONS](#)

[CE TABLES](#)

[CE LABSTAT DATABASE](#)

[CE PUBLIC USE MICRODATA](#)

[CE WORKSHOP AND SYMPOSIUM](#)

[CE GEOGRAPHIC DATA](#)

[CE EXPERIMENTAL RESEARCH PRODUCTS](#)

[CONTACT CE](#)

SEARCH CE

[GO](#)

CE TOPICS

[INFORMATION FOR CE RESPONDENTS](#)

[CE LIBRARY](#)

[CE METHODOLOGY](#)

CE Experimental Research Products

This page provides an introduction to special tabulation tables and other experimental products being developed by the Consumer Expenditure Survey Division. New research products will be added and updated periodically.

Information on this page includes the All Consumer Unit Prepublication Table (All CU Prepublication MVP), new cross-tabulated tables with higher incomes, the 2014 Higher Income Table which expanded the top income range to \$200,000 and over, a Generational Table that sorts expenditures by generation/cohorts, information on how to get the most from the CE published and prepublication tables, and the experimental state weight files.

Note that the 2014 through mid-2016 Generational tables were research work, and had not been produced using BLS production methods and standards. The 2016 annual Generational table is now included in the [standard tables](#).

On This Page

» [How to get the most from CE Tables](#)

» **NEW** [State Weights](#)

» [All CU Prepublication MVP Tables](#)

» **NEW** [Cross-Tabulated Tables](#)

» [Generational Research Tables](#)


» [2014 Higher Income Table](#)

<https://www.bls.gov/cex/csxresearchtables.htm>

Consumer Expenditure Surveys



SHARE ON:

PRINT: 

BROWSE CE

[CE HOME](#)[CE OVERVIEW](#)[CE NEWS RELEASES](#)[CE PUBLICATIONS](#)[CE TABLES](#)[CE LABSTAT DATABASE](#)[CE PUBLIC USE MICRODATA](#)[CE WORKSHOP AND SYMPOSIUM](#)[CE GEOGRAPHIC DATA](#)[CE EXPERIMENTAL RESEARCH PRODUCTS](#)[CONTACT CE](#)

SEARCH CE

Go

CE TOPICS

[INFORMATION FOR CE RESPONDENTS](#)

PUMD Documentation

This page contains documentation for the public-use microdata (PUMD) for years starting in 1996. Documentation for years prior to 1996 are available USB flash drive for [purchase \(PDF\)](#). The documentation falls into two major types: [Documentation that covers all years](#) and [documents that cover one particular year](#). If you are new to CE PUMD data, you may want to explore the [CE PUMD Getting Started Guide](#).

Documents covering all years

- [Consumer Expenditure Surveys Public-use Microdata Getting Started Guide](#) provides documentation for the CE PUMD, its conventions, files, sample code, and methodology.
- [Consumer Expenditure Surveys Program Considerations When Using the Public-use Microdata](#) discusses considerations when preparing and interpreting estimates with PUMD.
- [Dictionary for Interview and Diary Surveys \(XLSX\)](#) provides variables and codes from 1996 forward.
- [Source selection file \(XLSX\)](#) identifies which survey data variable comes from when combining the two CE surveys for 1996 forward.
- [Description of income imputation](#) provides information on the methods BLS uses to estimate income since 2004.

Documents covering specific years

- **Hierarchical groupings** lists the relation between the summary variables and their contributing variables as they are used in the published tables. Integrated stub (IntStub) lists the variables as BLS integrates them from the Interview and Diary Surveys.

https://www.bls.gov/cex/pumd_doc.htm

Documentation Covering All Years

- CE PUMD Getting Started Guide

- ▶ Provides documentation for the CE PUMD, its conventions, files, sample code, and methodology.

- CE Program Considerations When Using the Public-use Microdata

- ▶ Discusses considerations when preparing and interpreting estimates with PUMD

Documentation Covering All Years

- Dictionary for Interview and Diary Survey (XLSX)
 - ▶ Provides variables and codes from 1996 forward.
- Source selection file (XLSX)
 - ▶ Identifies which survey data item (category) comes from when combining the two CE surveys for 1996 forward.
- Description of Income Imputation
 - ▶ Provides information on the methods BLS uses to estimate income since 2004.

Documentation Covering Specific Years

- Hierarchical groupings
 - Establishes a relation between the published tables' series and their contributing expenditure classification codes (UCC).
- Sample code
 - Replicates totals in the CE published tables.
- Errata
 - Provides information on errors and their correction to the files of a particular year.

Consumer Expenditure Surveys



BROWSE CE

CE HOME

CE OVERVIEW

CE NEWS RELEASES

CE PUBLICATIONS

CE TABLES

CE LABSTAT DATABASE

CE PUBLIC USE MICRODATA

CE WORKSHOP AND SYMPOSIUM

CE GEOGRAPHIC DATA

CE EXPERIMENTAL RESEARCH PRODUCTS

CONTACT CE

SEARCH CE



CE TOPICS

PUMD Supplementary Information

CE provides the following supplementary resources about the PUMD.

- [CE Glossary](#)
- [CE FAQ](#)
- [User guide to income imputation \(PDF\)](#)
- [Protection of respondent confidentiality](#) provides information on the methods the CE survey uses to protect the confidentiality of its respondents.

Last Modified Date: August 3, 2017

RECOMMEND THIS PAGE USING:  Facebook  Twitter  LinkedIn

https://www.bls.gov/cex/pumd_sup.htm

Supplementary Resources

- [CE Glossary](#)
- [CE FAQ](#)
- [User's guide to income imputation \(PDF\)](#)
- [Protection of respondent confidentiality](#)



Consumer Expenditure Surveys



BROWSE CE

CE HOME
CE OVERVIEW
CE NEWS RELEASES
CE PUBLICATIONS
CE TABLES
CE LABSTAT DATABASE
CE PUBLIC USE MICRODATA®
CE WORKSHOP AND SYMPOSIUM
CE GEOGRAPHIC DATA
CE EXPERIMENTAL RESEARCH PRODUCTS
CONTACT CE

SEARCH CE



CE TOPICS

INFORMATION FOR CE RESPONDENTS
CE LIBRARY
CE METHODOLOGY
CE SURVEY MATERIALS
CE FAQ

Frequently Asked Questions (FAQs)

On this Page:

[Survey Overview](#)[Definitions](#)[Data Products](#)[Data Adjustments](#)[Survey Topics](#)[Survey Design](#)[Public Use Microdata \(PUMD\)](#)[Sampling and Nonsampling Errors](#)[\[Expand All\]](#)

Survey Overview

1. [What are the Consumer Expenditure Surveys?](#)
2. [How are the Consumer Expenditure Surveys used?](#)
3. [How do I contact the staff of the Consumer Expenditure Surveys?](#)
4. [How do the Census Bureau and BLS handle respondent confidentiality?](#)

Data Products

5. [What types of data are available and in what form?](#)
6. [What is the most recent Annual Report about the Consumer Expenditure Surveys data?](#)
7. [Are historical data from the Consumer Expenditure Surveys available?](#)
8. [How are Consumer Expenditure data used to estimate experimental poverty thresholds?](#)
9. [How are the data from the Interview and Diary Surveys integrated?](#)
10. [Do the published data come from both surveys?](#)

<https://www.bls.gov/cex/csxfqs.htm#PUMD>



2018 PUMD Release

■ September 10, 2019



Jimmy Choi

Economist

Division of Consumer Expenditure Surveys

202-691-7081

Choi.Jimmy@bls.gov

