

Consumer Finance Monthly – User Guide

Thank you for using the Consumer Finance Monthly (CFM) survey!

The purpose of this user guide is to providing guidelines in utilizing SCF. This user guide mainly covers issues related to unique survey properties and data analysis.

- Public vs. Private data

CFM dataset are available on-line to subscribing customers. Further information on how to subscribe to the data is also [available online](#). Almost all of the variables are available to the public. Some data are not released to the public due to privacy and its usefulness to subscribers. At times, interviewers would type in notes with regards to specific response. For instance, if a response to total income question is “\$50,000 to \$60,000” where an ideal response would be a single dollar amount, interviewers may enter \$55,000 and at the same time note that the respondent provided a range of income.

- Weights

Weights are used to make the statistics based on the sample in the survey comparable to the intended population, which is the adult population in the U.S. Weighting for CFM is stratified based on race and income variables, calibrated on age factor. Specifically, race is divided into Black (Only) and non-Black groups. Thereafter, non-Black group is divided into 11 income groups and Blacks into 2 income groups. These 13 groups are then compared to similar race-income household groups of Current Population Survey (CPS) of the same year as the CFM survey year. For 2006 weights, the CPS race-income stratification is based on [Table HINC-06](#), as shown in the Census website. Thereafter, these weights were raked based on their “performance” on household age factor based on [Table HINC-01](#) for 2006 weights.

In Stata, this weight is known as the analytical weight. This is important since Stata differentiates different types of weights.

- Missing data:

CFM treats missing data differently. Missing data types and their values are as follows:

1. Refused: -1
2. Don't know: -2
3. Invalid skip: -3
4. Valid missing: -4
5. Non-interview: -5

For a valid data that happen to have one of these five negative values (e.g. net worth), a “-10” would replace these values.

Web Investigator recognizes these values as missing data but the statistical programs that users utilize take these values as a valid, negative number. This issue is further discussed under the ‘Downloading the Data’ section.

- Brackets: Conventional vs. Unfolding

Wealth (assets and debt included), income, and their components’ variables may be obtained in three ways:

1. Exact value: An exact value is provided by respondents.
2. Conventional brackets: The mean of minimum and maximum values provided by respondents.
3. Unfolding brackets: A predetermined formula based on unfolding brackets.

The methods above are arranged according the methods presented to respondents to obtain the wealth, income, and their components’ variables.

Brackets are used to improve the estimations of variables that have “don’t know” or “refused” missing responses when respondents are asked for an exact estimation of their wealth, income, or their components’ values. Estimation of these types of missing observations ensures that missing observations are minimized.

Bracketing estimation technique is used on continuous variables in the CFM such as income, income’s components, and net worth’s components. There are two types of bracketing technique utilized in CFM: Conventional and unfolding. The conventional brackets are utilized first. If the conventional brackets result in “don’t know” or “refused” missing responses, the unfolding bracket technique is utilized.

With conventional bracketing, CFM obtains a minimum-maximum range in which people think or know their values lie on. The estimated value would be the average between the minimum and maximum values as indicated by the range.

With the unfolding bracket technique, a minimum-maximum range is determined by asking a series of question on whether the variable of interest is larger or smaller than various preset bracket values. Most questions in CFM have three preset values (e.g. \$100, \$2,000, \$10,000). All the variables in CFM that use bracketing technique cannot have values less than zero, thus, another preset value is \$0. For questions that have three preset values, there are 3 close brackets and 1 open bracket (e.g. (\$10,000, ∞)). The estimated values for the close brackets are the average of the minimum and maximum bracket values. The estimated value for the open bracket is the highest bracket value (\$10,000 in the example above), as pre-determined by CFM personnel.

- Created Variables (CVs)

CFM provides created variables (CVs) to make it easy to analyze CFM data. These CVs are based on respondent's actual response and the estimated value based on bracketing, both also available in the dataset. Income, net worth, and all other variables that make up income or net worth have CVs.

- Negative Asset?

Users might observe a few people with negative asset although arguably, asset and debt cannot be less than 0. Total non-financial or physical asset (CV-FINANCIAL) and thus, total asset (CV-ASSET) may take a negative number due to a very large deficit in business equity, defined as profit minus debt from business.

- Income does not include private transfer

For (best) total income that was derived based on summation of different types of income, the amount of private transfer is excluded from total income. The reasons for going so are twofold. One, private transfers can be very large that it might greatly overestimate a respondent's income. Two, private transfers should already be included in the wealth data. Including private transfer in both income and wealth overestimated people's financial well-being.

- Responses ending with "9"

For questions on income, wealth, and their components, dollar values that ends with a 9 (e.g. \$1,009, \$10,009) indicates that a pre-set maximum value for the particular variable has been reached. The respondent's exact response may be noted by the interviewer, which is available in the private dataset.

- Wave

Wave	Date	n	Weight Year
1	02/02/2005 – 09/12/2005	1,968	2005
2	09/12/2005 – 08/14/2006	1,943	2005 or 2006
3	06/06/2006 – 01/15/2007	3,364	2006
4	12/19/2006 – 08/31/2007	2,834	2007
5	09/01/2007 – 12/31/2007	877	2007
Total	02/02/2005 – 12/31/2007	10,986	2005 - 2007

Table 1: Wave Details

Weight Year	Date	n	Wave(s)
2005	02/02/2005 – 01/05/2006	3,282	1 & 2
2006	04/19/2006 – 01/15/2007	3,993	2 & 3
2007	12/19/2006 – 12/31/2007	3,711	4 & 5
Total	02/02/2005 – 12/31/2007	10,986	1 - 5

Table 2: Weight Year Details

Tables 1 and 2 provide an overall picture of the weight years and waves. A weight year entails the grouping of observations to be compared to intended population in order to derive sampling weights. Factors that determine the cutoffs of these weight years are the dates when respondents were interviewed and the survey wave. A survey wave groups together batches of respondents by interview time (about 3 to 9 months) and survey instrument version. Due to enrichments and improvements of the survey, CFM survey is constantly being updated.

- Downloading the Data

Before downloading the data, there are few things that might be useful to users:

1. Simple statistics.

Simple statistics can be derived from the extraction webpage. In fact, Web Investigator provides (unweighted) simple statistics when clicking on the variable name. For weighted simple statistics, after selecting the interested variable along with the appropriate weights, users may click on “Save/Download” and then, “Advanced Download Tab” tab (within the “Save/Download” page). The “Create Frequency / Table” with appropriate “Weight Option” selection allows users to arrive at weighted statistics.

2. Missing values.

When extracting data or running statistics from CHRR website, users have the option on whether users would like to leave missing variable coding “as is” or group them together as one category. If users choose to leave the coding of missing variables “as is”, the “Save/Download” in the Web Investigator would assign the appropriate labels to each missing value category, but statistical programs would take these ‘-1’, ‘-2’, ‘-3’, ‘-4’ and ‘-5’ responses as different types of missing values. For example, people who refused to provide their credit card interest rates are assigned ‘-2’, resulting in mean value of this variable to be underestimated. The default for missing value extraction is to group them all together as missing value or “.” in most statistical programs.

3. Important variable.

Don’t forget other important variables that might not directly be user’s variables of interest. Examples of these important variables are PUBLICID, weights, and weight year variables.

4. Extract all the variables.

Users with adequate computers are likely to be better off extracting all the whole completed dataset than to extract current interested variables. The advantage of doing this is that users do not have to go return to Web Investigator to extract other variables in instances where there have new interested variables or users forgot some other interested variables. Keep in mind that users might need to extract the dataset again if they are changes to the dataset in order to be up-to-date with the data.