Natalia LaSpada

Short Assignment 4

**Data Dictionary:**

**Entry 1:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** DUPERSID

**Description:** A unique ID variable for each person in the dataset, created by combining their five-digit, random, Dwelling Unit ID (DUID) with their person number (PID). The PID identifies each person within the dwelling unit. The values are assigned through the sampling process.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 60001101 - 60036528 | ID | 35,427 | 321,423,251 |

**Entry 2:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH31

**Description:** Perceived mental health status in Round 3 of Panel 19 or Round 1 of Panel 20. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 25 | 5,647 |
| -8 | Don’t know | 33 | 182,941 |
| -7 | Refused | 42 | 203,872 |
| -1 | Inapplicable | 1,038 | 4,120,335 |
| 1 | Excellent | 15,453 | 143,603,136 |
| 2 | Very good | 8,718 | 85,875,115 |
| 3 | Good | 7,753 | 67,019,926 |
| 4 | Fair | 1,887 | 16,020,317 |
| 5 | Poor | 478 | 4,391,962 |
| Total |  | 35,427 | 321,423,251 |

**Entry 3:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH42

**Description:** Perceived mental health status in Round 4 of Panel 19 or Round 2 of Panel 20. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 17 | 55,414 |
| -8 | Don’t know | 25 | 191,731 |
| -7 | Refused | 41 | 73,483 |
| -1 | Inapplicable | 586 | 3,329,760 |
| 1 | Excellent | 14,337 | 132,688,884 |
| 2 | Very good | 9,362 | 91,961,379 |
| 3 | Good | 8,494 | 71,113,266 |
| 4 | Fair | 2,059 | 17,421,274 |
| 5 | Poor | 506 | 4,588,060 |
| Total |  | 35,427 | 321,423,251 |

**Entry 4:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH53

**Description:** Perceived mental health status in Round 5 of Panel 19 or Round 3 of Panel 20. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 12 | 10,212 |
| -8 | Don’t know | 29 | 143,672 |
| -7 | Refused | 71 | 170,232 |
| -1 | Inapplicable | 316 | 3,678,546 |
| 1 | Excellent | 14,492 | 133,517,204 |
| 2 | Very good | 9,288 | 90,615,542 |
| 3 | Good | 8,723 | 73,017,722 |
| 4 | Fair | 2,016 | 16,155,456 |
| 5 | Poor | 480 | 4,114,664 |
| Total |  | 35,427 | 321,423,251 |

**Entry 5:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EDRECODE

**Description:** Education values, recoded, for all individuals. Values are grouped into broad, discrete categories determined based on responses to EDUCYR and HIDEG with EDUYRDG. Each individual receives a value (that is not -1) for either the original variables (EDUCYR and HIDEG) *or* for the variable EDUYRDG. Thus, both old and new education questions are combined. 327 individuals in Panel 19 had their education value edited based on EDUYRDG and corresponding National Health Interview Survey (NHIS) education data. The education value for Panel 20 individuals did not have to be edited in this manner.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 85 | 205,862 |
| -8 | Don’t know | 234 | 1,250,640 |
| -7 | Refused | 39 | 171,418 |
| -1 | Inapplicable | 3,050 | 24,900,178 |
| 1 | Less than or equal to 8th grade | 7,790 | 53,354,192 |
| 2 | 9th-12th grade with no high school degree or GED | 4,031 | 29,146,410 |
| 13 | GED or high school degree | 7,814 | 72,145,941 |
| 14 | Beyond high school, associate’s degree or college without a 4 year degree | 6,466 | 66,641,653 |
| 15 | 4-year college degree, bachelor’s degree | 3,766 | 45,744,405 |
| 16 | Master’s, doctorate, or professional degree | 2,152 | 27,862,552 |
| Total |  | 35,427 | 321,423,251 |

**Entry 6:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST31

**Description:** Employment status in Round 3 of Panel 19 or Round 1 of Panel 20. The values of EMPST31 represent individuals’ employment status prior to the imputation procedure used to create EMPST31H. Question was asked to all individuals 16 years of age or older. The following were mutually exclusive response options: “employed at RD 3/1 interview date” (1) indicated that the individual had employment at the interview date; “job to return to at RD 3/1 interview date” (2) indicated that the individual was not working during the reference period but had employment to return to as of the interview date; “job during RD 3/1 reference period” (3) indicated that the individual was working during the round but did not have employment at the interview date; and “not employed during RD 3/1” (4) indicated the individual was not working during the reference period, did not have employment at the interview date, and did not possess employment to return to.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 80 | 520,815 |
| -8 | Don’t know | 37 | 214,686 |
| -7 | Refused | 89 | 617,223 |
| -1 | Inapplicable | 9,522 | 69,404,213 |
| 1 | Employed at RD 3/1 Interview Date | 15,152 | 155,142,317 |
| 2 | Job to return to at RD 3/1 Interview Date | 107 | 946,328 |
| 3 | Job during RD 3/1 Reference Period | 316 | 3,143,413 |
| 4 | Not employed during RD 3/1 | 10,124 | 91,434,256 |
| Total |  | 35,427 | 321,423,251 |

**Entry 7:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST42

**Description:** Employment status in Round 4 of Panel 19 or Round 2 of Panel 20. The values of EMPST42 represent individuals’ employment status prior to the imputation procedure used to create EMPST42H. Question was asked to all individuals 16 years of age or older. The following were mutually exclusive response options: “employed at RD 4/2 interview date” (1) indicated that the individual had employment at the interview date; “job to return to at RD 4/2 interview date” (2) indicated that the individual was not working during the reference period but had employment to return to as of the interview date; “job during RD 4/2 reference period” (3) indicated that the individual was working during the round but did not have employment at the interview date; and “not employed during RD 4/2” (4) indicated the individual was not working during the reference period, did not have employment at the interview date, and did not possess employment to return to.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 37 | 235,479 |
| -8 | Don’t know | 22 | 94,308 |
| -7 | Refused | 72 | 228,005 |
| -1 | Inapplicable | 9,075 | 68,138,638 |
| 1 | Employed at RD 4/2 Interview Date | 15,617 | 157,347,856 |
| 2 | Job to return to at RD 4/2 Interview Date | 28 | 326,434 |
| 3 | Job during RD 4/2 Reference Period | 1,040 | 9,697,823 |
| 4 | Not employed during RD 4/2 | 9,536 | 85,354,707 |
| Total |  | 35,427 | 321,423,251 |

**Entry 8:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST53

**Description:** Employment status in Round 5 of Panel 19 or Round 3 of Panel 20. The values of EMPST53 represent individuals’ employment status prior to the imputation procedure used to create EMPST53H. Question was asked to all individuals 16 years of age or older. The following were mutually exclusive response options: “employed at RD 5/3 interview date” (1) indicated that the individual had employment at the interview date; “job to return to at RD 5/3 interview date” (2) indicated that the individual was not working during the reference period but had employment to return to as of the interview date; “job during RD 5/3 reference period” (3) indicated that the individual was working during the round but did not have employment at the interview date; and “not employed during RD 5/3” (4) indicated the individual was not working during the reference period, did not have employment at the interview date, and did not possess employment to return to.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -9 | Not ascertained | 27 | 124,051 |
| -8 | Don’t know | 24 | 116,188 |
| -7 | Refused | 70 | 392,454 |
| -1 | Inapplicable | 8,736 | 67,518,586 |
| 1 | Employed at RD 5/3 Interview Date | 15,741 | 157,253,618 |
| 2 | Job to return to at RD 5/3 Interview Date | 59 | 542,493 |
| 3 | Job during RD 5/3 Reference Period | 1,030 | 8,550,171 |
| 4 | Not employed during RD 5/3 | 9,740 | 86,925,690 |
| Total |  | 35,427 | 321,423,251 |

**Entry 9:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** POVCAT15

**Description:** Categorical values for family income as percent of poverty line. The categories are created by breaking the continuous income values into the intervals that were designated by the MEPS HC-181 Codebook.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 1 | Poor / negative | 7,924 | 43,502,893 |
| 2 | Near poor | 2,260 | 14,177,868 |
| 3 | Low income | 5,921 | 44,382,151 |
| 4 | Middle income | 10,111 | 91,924,509 |
| 5 | High income | 9,211 | 127,435,829 |
| Total |  | 35,427 | 321,423,251 |

**Entry 10:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** TTLP15X

**Description:** Continuous values for individual person’s total income. Represents the total, adjusted person-level income, which was created by adding all income components together, as referenced in Entry 8 to correspond with the definition of total person income that most closely corresponds with the Current Population Survey (CPS) definition of income. There are no missing value for person-level income because all missing values were estimated by using logical editing and weighted, sequential hot-decks.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -$149,923 - $409,118 | Total individual income | 35,427 | 321,423,251 |

**Entry 11:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** FAMINC15

**Description:** Continuous values for each CPS family’s total income. Family total income was calculated by summing person-level income totals for each family member. There are no missing value for family-level income because all missing values were estimated by using logical editing and weighted, sequential hot-decks.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -$134,723 - $521,685 | Total family income | 35,427 | 321,423,251 |

**Entry 12:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST31H

**Description:** Employment status in Round 3 of Panel 19 or Round 1 of Panel 20. All individuals have a value for EMPST31H because the missing values from EMPST31 are imputed, using weighted sequential hot-decking, to determine their value in EMPST31H. Therefore, the categories of EMPST31 that were imputed were the following: -9 (not ascertained), -8 (don’t know), and -7 (refused). Additionally, EMPST31 categories 3 (employed during the reference period) and 4 (not employed with no job to return to) were combined into the single category, 34 (not employed during RD 3/1). All individuals with EMPST31H value of 1 or 2 have valid information regarding their current employment.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -1 | Inapplicable | 9,522 | 69,404,213 |
| 1 | Employed at RD 3/1 Int Date | 15,286 | 155,933,622 |
| 2 | Job to Return to at RD 3/1 Int Date | 107 | 946,328 |
| 34 | Not Employed During RD 3/1 | 10,512 | 95,139,088 |
| Total |  | 35,427 | 321,423,251 |

**Entry 13:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST42H

**Description:** Employment status in Round 4 of Panel 19 or Round 2 of Panel 20. All individuals have a value for EMPST42H because the missing values from EMPST42 are imputed, using weighted sequential hot-decking, to determine their value in EMPST42H. Therefore, the categories of EMPST42 that were imputed were the following: -9 (not ascertained), -8 (don’t know), and -7 (refused). Additionally, EMPST42 categories 3 (employed during the reference period) and 4 (not employed with no job to return to) were combined into the single category, 34 (not employed during RD 4/2). All individuals with EMPST42H value of 1 or 2 have valid information regarding their current employment.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -1 | Inapplicable | 9,075 | 68,138,638 |
| 1 | Employed at RD 4/2 Int Date | 15,695 | 157,610,611 |
| 2 | Job to Return to at RD 4/2 Int Date | 28 | 326,434 |
| 34 | Not Employed During RD 4/2 | 10,629 | 95,347,568 |
| Total |  | 35,427 | 321,423,251 |

**Entry 14:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST53H

**Description:** Employment status in Round 5 of Panel 19 or Round 3 of Panel 20. All individuals have a value for EMPST53H because the missing values from EMPST53 are imputed, using weighted sequential hot-decking, to determine their value in EMPST53H. Therefore, the categories of EMPST53 that were imputed were the following: -9 (not ascertained), -8 (don’t know), and -7 (refused). Additionally, EMPST53 categories 3 (employed during the reference period) and 4 (not employed with no job to return to) were combined into the single category, 34 (not employed during RD 5/3). All individuals with EMPST53H value of 1 or 2 have valid information regarding their current employment.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -1 | Inapplicable | 8,736 | 67,518,586 |
| 1 | Employed at RD 5/3 Int Date | 15,821 | 157,653,058 |
| 2 | Job to Return to at RD 5/3 Int Date | 59 | 542,493 |
| 34 | Not Employed During RD 5/3 | 10,811 | 95,709,114 |
| Total |  | 35,427 | 321,423,251 |

**Entry 15:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH31\_recode

**Description:** Perceived mental health status, recoded, in Round 3 of Panel 19 or Round 1 of Panel 20. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 100 | 392,460 |
| -1 | Inapplicable | 1,038 | 4,120,335 |
| 1 | Excellent | 15,453 | 143,603,136 |
| 2 | Very good | 8,718 | 85,875,115 |
| 3 | Good | 7,753 | 67,019,926 |
| 4 | Fair | 1,887 | 16,020,317 |
| 5 | Poor | 478 | 4,391,962 |
| Total |  | 35,427 | 321,423,251 |

**Entry 16:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH42\_recode

**Description:** Perceived mental health status, recoded, in Round 4 of Panel 19 or Round 2 of Panel 20. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 83 | 320,628 |
| -1 | Inapplicable | 586 | 3,329,760 |
| 1 | Excellent | 14,337 | 132,688,884 |
| 2 | Very good | 9,362 | 91,961,379 |
| 3 | Good | 8,494 | 71,113,266 |
| 4 | Fair | 2,059 | 17,421,274 |
| 5 | Poor | 506 | 4,588,060 |
| Total |  | 35,427 | 321,423,251 |

**Entry 17:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH53\_recode

**Description:** Perceived mental health status, recoded, in Round 5 of Panel 19 or Round 3 of Panel 20. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 112 | 324,116 |
| -1 | Inapplicable | 316 | 3,678,546 |
| 1 | Excellent | 14,492 | 133,517,204 |
| 2 | Very good | 9,288 | 90,615,542 |
| 3 | Good | 8,723 | 73,017,722 |
| 4 | Fair | 2,016 | 16,155,456 |
| 5 | Poor | 480 | 4,114,664 |
| Total |  | 35,427 | 321,423,251 |

**Entry 18:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EDRECODE\_recode

**Description:** Recoded education values, recoded again to group all missing data together, for all individuals. Values are grouped into broad, discrete categories determined based on responses to EDUCYR and HIDEG with EDUYRDG. Each individual receives a value (that is not -1) for either the original variables (EDUCYR and HIDEG) *or* for the variable EDUYRDG. Thus, both old and new education questions are combined. 327 individuals in Panel 19 had their education value edited based on EDUYRDG and corresponding National Health Interview Survey (NHIS) education data. The education value for Panel 20 individuals did not have to be edited in this manner. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 358 | 1,627,920 |
| -1 | Inapplicable | 3,050 | 24,900,178 |
| 1 | Less than or equal to 8th grade | 7,790 | 53,354,192 |
| 2 | 9th-12th grade with no high school degree or GED | 4,031 | 29,146,410 |
| 13 | GED or high school degree | 7,814 | 72,145,941 |
| 14 | Beyond high school, associate’s degree or college without a 4 year degree | 6,466 | 66,641,653 |
| 15 | 4-year college degree, bachelor’s degree | 3,766 | 45,744,405 |
| 16 | Master’s, doctorate, or professional degree | 2,152 | 27,862,552 |
| Total |  | 35,427 | 321,423,251 |

**Entry 19:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST31\_recode

**Description:** Employment status, recoded, in Round 3 of Panel 19 or Round 1 of Panel 20. The values of EMPST31 represent individuals’ employment status prior to the imputation procedure used to create EMPST31H. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing. Question was asked to all individuals 16 years of age or older.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 206 | 1,352,724 |
| -1 | Inapplicable | 9,522 | 69,404,213 |
| 1 | Employed at RD 3/1 Interview Date | 15,152 | 155,142,317 |
| 2 | Job to return to at RD 3/1 Interview Date | 107 | 946,328 |
| 3 | Job during RD 3/1 Reference Period | 316 | 3,143,413 |
| 4 | Not employed during RD 3/1 | 10,124 | 91,434,256 |
| Total |  | 35,427 | 321,423,251 |

**Entry 20:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST42\_recoded

**Description:** Employment status, recoded, in Round 4 of Panel 19 or Round 2 of Panel 20. The values of EMPST42 represent individuals’ employment status prior to the imputation procedure used to create EMPST42H. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing. Question was asked to all individuals 16 years of age or older.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 131 | 557,792 |
| -1 | Inapplicable | 9,075 | 68,138,638 |
| 1 | Employed at RD 4/2 Interview Date | 15,617 | 157,347,856 |
| 2 | Job to return to at RD 4/2 Interview Date | 28 | 326,434 |
| 3 | Job during RD 4/2 Reference Period | 1,040 | 9,697,823 |
| 4 | Not employed during RD 4/2 | 9,536 | 85,354,707 |
| Total |  | 35,427 | 321,423,251 |

**Entry 21:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST53\_recoded

**Description:** Employment status, recoded, in Round 5 of Panel 19 or Round 3 of Panel 20. The values of EMPST53 represent individuals’ employment status prior to the imputation procedure used to create EMPST53H. Recoding involved indicating all missing data with a new value of -6. Values of -9 (not ascertained), -8 (don’t know), and -7 (refused) were determined to be “missing” and therefore were recoded as -6. Values of -1 (inapplicable), however, were not recoded because they were not considered to be missing. Question was asked to all individuals 16 years of age or older.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| -6 | Missing | 121 | 632,693 |
| -1 | Inapplicable | 8,736 | 67,518,586 |
| 1 | Employed at RD 5/3 Interview Date | 15,741 | 157,253,618 |
| 2 | Job to return to at RD 5/3 Interview Date | 59 | 542,493 |
| 3 | Job during RD 5/3 Reference Period | 1,030 | 8,550,171 |
| 4 | Not employed during RD 5/3 | 9,740 | 86,925,690 |
| Total |  | 35,427 | 321,423,251 |

**Entry 22:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH31\_recode3

**Description:** Perceived mental health status in Round 3 of Panel 19 or Round 1 of Panel 20, recoded as dichotomous (missing or not missing). Recoding involved representing all missing data with a new value of 1, and representing all non-missing data (including inapplicable values) with a value of 0. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family. The second recoding (MNHLTH31\_recode2) was used as a step between the initial recoding (MNHLTH31\_recode) and the dichotomous recode (MNHLTH31\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Non-missing | 35,327 | 321,030,791 |
| 1 | Missing | 100 | 392,460 |
| Total |  | 35,427 | 321,423,251 |

**Entry 23:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH42\_recode3

**Description:** Perceived mental health status in Round 4 of Panel 19 or Round 2 of Panel 20, recoded as dichotomous (missing or not missing). Recoding involved representing all missing data with a new value of 1, and representing all non-missing data (including inapplicable values) with a value of 0. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family. The second recoding (MNHLTH42\_recode2) was used as a step between the initial recoding (MNHLTH42\_recode) and the dichotomous recode (MNHLTH42\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Non-missing | 35,344 | 321,102,623 |
| 1 | Missing | 83 | 320,628 |
| Total |  | 35,427 | 321,423,251 |

**Entry 24:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** MNHLTH53\_recode3

**Description:** Perceived mental health status in Round 5 of Panel 19 or Round 3 of Panel 20, recoded as dichotomous (missing or not missing). Recoding involved representing all missing data with a new value of 1, and representing all non-missing data (including inapplicable values) with a value of 0. Question was asked to all current or institutionalized persons of any age. Persons rated each member in the family. The second recoding (MNHLTH53\_recode2) was used as a step between the initial recoding (MNHLTH53\_recode) and the dichotomous recode (MNHLTH53\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Non-missing | 35,315 | 321,099,135 |
| 1 | Missing | 112 | 324,116 |
| Total |  | 35,427 | 321,423,251 |

**Entry 25:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EDRECODE\_recode3

**Description:** Education values that are given to each individual and then recoded as either missing (1) or non-missing (0). To give each individual an education value, old and new education questions were combined and individuals were given an education value for either the original variables (EDUCYR and HIDEG) *or* for the variable EDUYRDG. These recoded education values were recoded again to group all missing data together (EDRECODE\_recode) and then recoded *again* into dichotomous missing (1) or non-missing (including inapplicable values) (0) data for all individuals (EDRECODE\_recode3). The second recoding (EDRECODE\_recode2) was used as a step between the initial recoding (EDRECODE\_recode) and the dichotomous recode (EDRECODE\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Non-missing | 35,069 | 319,795,331 |
| 1 | Missing | 358 | 1,627,920 |
| Total |  | 35,427 | 321,423,251 |

**Entry 26:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST31\_recode3

**Description:** Employment status in Round 3 of Panel 19 or Round 1 of Panel 20, recoded as dichtomous (missing or not missing). Recoding involved representing all missing data with a new value of 1, and representing all non-missing data (including inapplicable values) with a value of 0. Recoding 1 (EMPST31\_recode) involved grouping all missing data as -6, and recoding 2 (EMPST31\_recode2) was used as a step to create the dichotomous recode (EMPST31\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Not missing | 35,221 | 320,070,527 |
| 1 | Missing | 206 | 1,352,724 |
| Total |  | 35,427 | 321,423,251 |

**Entry 27:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST42\_recode3

**Description:** Employment status in Round 4 of Panel 19 or Round 2 of Panel 20, recoded as dichtomous (missing or not missing). Recoding involved representing all missing data with a new value of 1, and representing all non-missing data (including inapplicable values) with a value of 0. Recoding 1 (EMPST42\_recode) involved grouping all missing data as -6, and recoding 2 (EMPST42\_recode2) was used as a step to create the dichotomous recode (EMPST42\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Not missing | 35,296 | 320,865,459 |
| 1 | Missing | 131 | 557,792 |
| Total |  | 35,427 | 321,423,251 |

**Entry 28:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** EMPST53\_recode3

**Description:** Employment status in Round 5 of Panel 19 or Round 3 of Panel 20, recoded as dichtomous (missing or not missing). Recoding involved representing all missing data with a new value of 1, and representing all non-missing data (including inapplicable values) with a value of 0. Recoding 1 (EMPST53\_recode) involved grouping all missing data as -6, and recoding 2 (EMPST53\_recode2) was used as a step to create the dichotomous recode (EMPST53\_recode3). The dichotomous recode was used to evaluate the type of missing data within the variable, by testing for independence between the various recoded dichotomous variables.

**Data type/format:** numeric

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| 0 | Not missing | 35,306 | 320,790,558 |
| 1 | Missing | 121 | 632,693 |
| Total |  | 35,427 | 321,423,251 |

**New Variables Created:**

**Entry 29:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** TTLP15X\_cat1

**Description:** A newly created variable with categorical values for individual person’s total income, created from TTLP15X, which is the continuous version of total person income. The categories are created by breaking the continuous income values into the intervals that were designated by the MEPS HC-181 Codebook. Total, adjusted person-level income corresponds to the continuous variable TTLP15X, which was created by adding all income components together (“annual earnings from wages, salaries, bonuses, tips, commissions; business and farm gains and losses; unemployment and workers’ compensation; interest and dividends; alimony, child support, and other private cash transfers; private pensions, IRA withdrawals, social security, and veterans payments; supplemental security income and cash welfare payments from public assistance, Temporary Assistance for Needy Families, and related programs; gain or losses from estates, trusts, partnerships, S corporations, rent, and royalties; and a small amount of “other” income) except SALEP15X (tax refunds and capital gains), because this definition of total person income most closely corresponds with the Current Population Survey (CPS) definition of income. There are no missing value for person-level income because all missing values were estimated by using logical editing and weighted, sequential hot-decks.

**Data type/format:** categorical

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| Negative | -$149,923- -$58 | 11 | 109,985 |
| None | 0.00 | 12,068 | 89,996,708 |
| Very low | $2 - $12,000 | 5,841 | 46,057,754 |
| Low | $12,001 - $24,082 | 5,833 | 49,941,506 |
| Medium | $24,083 - $45,000 | 5,864 | 59,156,447 |
| High | $45,001 - $409,118 | 5,810 | 76,160,852 |
| Total |  | 35,427 | 321,423,251 |

**Entry 30:**

**Dataset:** MEPS HC181, 2015 Full year consolidated codebook

**R Variable Name:** FAMINC15\_cat1

**Description:** A newly created variable with categorical values for each CPS family’s total income, created from FAMINC15, which is the continuous version of total family income. The categories are created by breaking the continuous income values into the intervals that were designated by the MEPS HC-181 Codebook. Family total income was calculated by summing person-level income totals for each family member. There are no missing value for family-level income because all missing values were estimated by using logical editing and weighted, sequential hot-decks.

**Data type/format:** categorical

**Values and frequencies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Value** | **Description** | **Unweighted Frequency** | **Weighted Frequency** |
| Negative | -$134,723 - -$523 | 9 | 48,748 |
| None | 0.00 | 1,161 | 6,246,944 |
| Very low | $9 - $23,000 | 8,593 | 55,616,975 |
| Low | $23,001 - $47,000 | 8,563 | 66,442,676 |
| Medium | $47,001 - $85,514 | 8,537 | 82,757,829 |
| High | $85,515 - $521,685 | 8,564 | 110,310,080 |
| Total |  | 35,427 | 321,423,251 |