**Example**

**PSC-4099**

**Codebook Assignment**

**Research Question:** Are wealthier Americans more supportive of gay marriage?

**Dataset:** 2016 CCES

**Dependent Variable(s)**

|  |  |  |
| --- | --- | --- |
| **Original Question** | **Original Coding** | **Your Coding (Variable type)** |
| CC16\_335  Do you favor or oppose allowing gays and lesbians to marry legally? | 1 Favor  2 Oppose  8 skipped  9 not asked | Favorgaymar (dummy)  2🡪 0=Oppose  1🡪 1=Favor  8, 9, missing 🡪 missing |

**Independent Variable(s)**

|  |  |  |
| --- | --- | --- |
| **Age** | **Original Coding** | **Your Coding** |
| age | Original values range from 18 to 96.  97 97 years or older  98 Prefer not to say  \*Note, when there is a long list choices, you don’t need to list every value; the point here is (1) to make it clear what needs to be recoded and (2) to make sure that \*every\* value is accounted for in the recoding.   For example, if you are coding lots of religious denominations into a handful of categories, you might have something like this for what is in the original survey rather than writing every group out (I am making these up) 1 = Catholic,  2 = Evan Prot.  3 thu 7, 11, 14 = Various mainline Protestant groups;  8,9,10,12,13 = Other religions that will be grouped into an “other category”;  15 thru 17 = various types of non-religious people.  90= DK or non-response.  If you do not list every value for an original variable, double check very closely to make sure that every response category is covered and correct. | **AgeInYears** (interval)  18 thru 97🡪 same values;  98, missing 🡪 missing  **Under 50** (dummy)  18 thru 40🡪 1=Under 50 yrs;  50 thru 97 🡪 0=Not under 50  98, missing 🡪 missing |

**\***

|  |  |  |
| --- | --- | --- |
| **Original Question** | **Original Coding** | **Your Coding** |
| faminc  Thinking back over the last year, what was your family's annual income? | 1 Less than $10,000  2 $10,000 - $19,999  3 $20,000 - $29,999  4 $30,000 - $39,999  5 $40,000 - $49,999  6 $50,000 - $59,999  7 $60,000 - $69,999  8 $70,000 - $79,999  9 $80,000 - $99,999  10 $100,000 - $119,999  11 $120,000 - $149,999  12 $150,000 - $199,999  13 $200,000 - $249,999  14 $250,000 - $349,999  15 $350,000 - $499,999  16 $500,000 or more  97 Prefer not to say | **incless50k** (dummy)  6 thru 16🡪 0=Over 50k;  1 thru 5 🡪 1=50k or under;  97, missing 🡪 missing  **inc50\_100k** (dummy)  1 thru 5, 10 thru 16🡪 0=Not between 50k-100k  6 thru 9 🡪 1=Between 50k-100k  97, missing 🡪 missing  **inc100\_150k** (dummy)  1 thru 16🡪 0=Not between 100k-150k  10, 11 🡪 1=Between 50k-100k  97, missing 🡪 missing  **incgreater150k**  1 thru 11🡪 0=Less than 150k  12 thru 16 🡪 1=150k or greater  97, missing 🡪 missing |

**Control Variables**

|  |  |  |
| --- | --- | --- |
| **Original Question** | **Original Coding** | **Your Coding** |
| gender  Are you male or female? | 1 Male  2 Female  8 skipped  9 not asked | Female (dummy)  1 🡪 0=Male  2 🡪 1=Female  8, 9, missing 🡪 missing |
| education  What is the highest level of education you have completed? | 1 No HS  2 High school graduate  3 Some college  4 2-year  5 4-year  6 Post-grad  8 skipped  9 not asked | collegegrad  1, 2, 3, 4 🡪 0=Not college grad  5, 6 🡪 1=College or higher  8, 9, missing 🡪 missing |
| ideol  Which of the words below best describes your political ideology? | 1 Very liberal  2 Liberal  3 Somewhat liberal  4 Neither liberal, nor conservative  5 Somewhat conservative  6 Conservative  7 Very Conservative  8 skipped  9 not asked | LiberalIndex0to1 (continuous, 0-1 rescaled)   1. 8, 9, missing 🡪 missing 2. 1 thru 7 reverse coded to create cons -> liberal scale (i.e. 8 – original value) 3. (Original value)/7 to create index ranging from 0 to 1. |
| Note:  I have only shown two examples here, but you should show your full battery of control variables. |  |  |