

(Abridged) Codebook for the March 2008 Current Population Survey

The Current Population Survey is a monthly survey carried out by the Bureau of Labor Statistics. It is used to compute national labor force statistics, including the unemployment rate. With about a one-year lag, they release the individual data records for the survey participants themselves. I have included for you the survey from March of last year.

I selected a subset of variables and observations to work with, but these are basically the data in their raw form supplied by the government. Many of these variables are coded categorical variables. For example occupation is coded by several numbers, each of which corresponds to a different occupation. Without a “key” – called the “codebook” – to decode these variables, these numbers are meaningless. For example, you would never want to put the “occupation” variable directly into the regression; instead, you would first create a series of dummy variables for all the different occupations. I will show you how to do that.

Part of the purpose of this is to show you that data are often not “pretty” and require some cleaning up before doing useful research. I will show you how to do this in class. For your reference, below is a list of all the variables included in the data. I leave it mostly in its original ugly form, but I will go over some of the pieces of it with you

D <u>HHID</u> 5 2 (00001:99999)	D <u>RACE</u> 2 24 (01:21)
Household sequence number	Race
V All households	U All
V 00001- .Household sequence	V 01 .White only
number	V 02 .Black only
V 99999 .	V 03 .American Indian,
	V .Alaskan Native only (AI)
	V 04 .Asian only
	V 05 .Hawaiin/Pacific Islander
	V .only (HP)
D <u>REGION</u> 1 39 (1:4)	
Region	
U All households	
V 1 .Northeast	
V 2 .Midwest	
V 3 .South	
V 4 .West	
D <u>FIPST</u> 2 42 (01:56)	D <u>FNLWGT</u> 8 50
State FIPS code	(00000000:99999999)
V 01-56 .State code	Final weight (2 implied decimal
	places)
D <u>CBSA</u> 5 44 (00000:79600)	V 00000000 .Supplemental Spanish
Metropolitan CBSA FIPS CODE	sample
V 0000 .Non-met or not	V 00000001 - .March basic sample
identified	weight
V 00460- .CBSA code	V 99999999 .
V 79600 .	
D <u>SEX</u> 1 20 (1:2)	D <u>USLHRS</u> 2 128 (00:99)
Item 18g - Sex	How many hrs per week does ...
U All	usually work at this job?
V 1 .Male	U All
V 2 .Female	V -4 .Hours vary
	V -1 .Not in universe
	V 00 .None, no hours
	V 01-99 .Entry

D HRLYWK 1 130 (0:2)
Is ... paid by the hour on this job?

V 0 .Not in universe or children and
V .Armed Forces
V 1 .Yes
V 2 .No

D UNION 1 140 (0:2)
On this job, is ... covered by a union

or employee association contract
U A-UNMEM=2
V 0 .Not in universe or children and
V .Armed Forces
V 1 .Yes
V 2 .No

D LFSTAT 1 145 (0:7)
Labor force status recode

U All
V 0 .Children or Armed Forces
1 .Working
2 .With job, not at work
V 3 .Unemployed, looking for work
V 4 .Unemployed, on layoff
V 7 .NILF**

D MAJIND 2 155 (00:14)
Major industry code

U A-CLSWKR = 1-7
V 0 .Not in universe, or children
V 1 .Agriculture, forestry, fishing, and hunting
V 2 .Mining
V 3 .Construction
V 4 .Manufacturing
V 5 .Wholesale and retail trade
V 6 .Transportation and utilities
V 7 .Information
V 8 .Financial activities
V 9 .Professional and business services
V 10 .Educational and health services
V 11 .Leisure and hospitality
V 12 .Other services
V 13 .Public administration
V 14 .Armed Forces

D MAJOCC 2 159 (00:11)
Major occupation recode

U A-CLSWKR = 1-7
V 0 .Not in universe or children
V 1 .Management, business, and financial occupations
V 2 .Professional and related occupations
V 3 .Service occupations
V 4 .Sales and related occupations
V 5 .Office and administrative support occupations
V 6 .Farming, fishing, and forestry occupations
V 7 .Construction and extraction occupations
V 8 .Installation, maintenance, and repair occupations
V 9 .Production occupations
V 10 .Transportation and material moving occupations
V 11 .Armed Forces

D WKSWORK 2 171 (00:52)
Item 33 - During 20.. in how many weeks

did ... work even for a few hours include paid vacation and sick leave as work.
U WORKYN = 1
V 00 .Not in universe
V 01 .1 week
V ...
V 52 .52 weeks

D HEALTHINS 1 472 (0:2)
Item 75b - Was ... covered by private

health insurance plan
U HHI-YN = 1
V 0 .Not in universe
V 1 .Yes
V 2 .No

**NILF stands for "not in the labor force"

The other variables in the dataset I constructed. They are:

wage hr - hourly wage

hisp - a dummy variable equal to 1 for Hispanics, zero otherwise

yrsed - (approximate) years of education completed, based on degree