

To: ANES file
 From: Matthew DeBell
 Date: October 13, 2014, revised March 16, 2015
 Re: Plan to update race & ethnicity variables on the ANES CDF

This memo describes the race and ethnicity variables on ANES Time Series, describes their summaries on the ANES Cumulative Data File, discusses difficulties with these variables, and proposes versions of these variables for the next CDF update. Proposals are on page 7.

The memo follows the Census Bureau approach of treating race and ethnicity as two concepts. A person's ethnic identity as Hispanic or not Hispanic is independent of race; Hispanics may be of any race. This is an organizing principle for the proposals here.

Race & Ethnicity Variables in the Time Series

Race/ethnicity has been measured in different ways over time. Instrumentation and response categories are summarized below for each Time Series study year.

Race and race/ethnicity instrumentation and response categories, by study year

<i>Years</i>	<i>Instrumentation</i>	<i>Response Categories</i>
1948, 1952, 1956, 1958, 1960, 1962, 1964	Interviewer observation: Race of R	White Negro Other, including Mexicans, Puerto Ricans, Asians (or "Other" in 1948, '52) NA (not 1960)
1966, 1968, 1970, 1972, 1974, 1976	Interviewer observation: Respondent's race is:	White Negro ("Black" in 1974, 1976) Puerto Rican Mexican-American, Chicano ("Mexican" in '66, '68) Oriental American Indian Other (not 1966) NA
1978, 1980, 1982, 1984, 1986, 1988, 1990, 1992, 1994, 1996, 1998	Interviewer observation: Respondent's race is:	White Black American Indian or Alaska Native Asian or Pacific Islander Other [not all years; does not include Hispanics] NA
2000, 2002, 2004, 2008	What racial or ethnic group or groups best describes you? [multiple mentions recorded]	Black Asian Native American Hispanic or Latino (just Hispanic in 2004) White Other
2012	FtF only: I am going to read you a list of five race categories. Both FtF and web: Please choose one or more races that you consider yourself to be. [multiple mentions recorded]	White Black or African-American American Indian or Alaska Native Asian Native Hawaiian or other Pacific Islander Other

Ethnicity instrumentation and response categories, by study year

<i>Years</i>	<i>Instrumentation</i>	<i>Response Categories</i>
1978, 1980, 1982, 1984, 1986	Interviewer observation: Is R of Hispanic origin?	Yes, Mexican-American, Chicano Yes, Puerto Rican Yes, other Hispanic No
1972, 1974, 1976, 1980, 1982, 1984, 1986, 1988, 1990, 1992, 1994, 1996, 1998, 2000, 2004, 2008, 2012	In addition to being American, what do you consider your main ethnic group or nationality group?	[multiple codes for nationalities and groups including German, Chinese, Mexican, Mexican-American, Jewish, etc]
1978	In addition to being American, is there another nationality or ethnic group that you feel you belong to? (IF YES:) What group is that?	
1988, 1990, 1992, 1994, 1996, 1998, 2000, 2002, 2004, 2008	[if Hispanic origin not mentioned above] Are you of Spanish or Hispanic origin or descent?	Yes No
1988, 1990, 1992, 1994, 1996, 1998, 2000, 2002, 2004, 2008	Please look at (page X of) the booklet and tell me which category best describes your Hispanic origin. 2002: Would you describe your Hispanic origin as... [one response only]	Mexican Puerto Rican Cuban Latin American Central American Spanish Other
2012	Are you Spanish, Hispanic, or Latino?	Yes No
2012	Which Hispanic group are you? [multiple mentions recorded]	Mexican Mexican-American Chicano Puerto Rican Cuban Cuban-American Other
2012	Which group do you most closely identify with?	Mexican Mexican-American Chicano Puerto Rican Cuban Cuban-American Other

Note: Other versions existed for telephone cases. For example in 1984 a question was, "Would you mind telling me your race? Are you white, black, American Indian or Alaska Native, Asian or Pacific Islander?" These matched the coding categories in use for face-to-face cases at the time.

Race and Ethnicity Variables on the Cumulative Data File

Race and ethnicity are summarized in six variables on the CDF. The current summaries stop in 2004; there are no race/ethnicity data on the current cumulative file for 2008.

These six variables are summarized below and are fully described in the Appendix.

2-category Race, VCF0105

- White
- Black
- Other/missing

3-category Race, VCF0106

- White
- Black
- Other
- Missing

6-category Race, VCF0106a

- White
- Black
- Asian
- Native American
- Hispanic
- Other
- Missing

Hispanic Origin Type, VCF0107

- Mexican-American, Chicano
- Puerto Rican
- Other Hispanic
- Hispanic but DK/NA type
- Not Hispanic

Hispanic Origin, VCF0108

- Hispanic
- Not Hispanic
- Missing [multiple codes]

Ethnicity, VCF0109

- Dozens of codes for national origin.

Difficulties with Current Variables

There are characteristics of the Time Series variables and with the summary variables on the CDF that present difficulties for data processing and data analysis.

Difficulties with the Time Series

There have been significant changes in the measurement of race and of ethnicity over time.

From 1948 through 1976, the single race/ethnicity variable records one response in a set of exhaustive and mutually exclusive race/ethnicity categories, so we have no data on race for Hispanics. From 1978 forward, race is recorded separately from ethnicity, so we apparently have race data for Hispanics.

The measurement of Hispanic origins has varied over time. Distinctions among Mexican-American, Puerto Rican, and other origins were not available before 1972. They have always been available since 1972, but the distinctions have not been measured using consistent methods. From 1972 to 1976 Hispanic origins were ascertained with an open-ended question with no probing for distinctions among particular Hispanic origins. From 1978-1986 there was an interviewer observation of detailed Hispanic origin, but we have no information on the particular methods interviewers used to discern this information or whether they may have unreliably guessed. From 1988 through 2008, respondents were asked if they are “of Spanish or Hispanic origin or descent,” with a follow-up question for detailed origins, and in 2012 they were asked substantively the same information with different wording and response categories.

The propensity of Hispanic respondents to identify as “white” is strongly dependent on question wording and question order. The Time Series has changed the wording and order of questions measuring Hispanic identity and race, including whether respondents are asked if they are Hispanic before or after they are asked their race, so the Time Series data on race for Hispanic respondents may exhibit substantial measurement error due to questionnaire effects. In 2000-2008 race is missing for many Hispanics because the question was worded to make it seem satisfactory for Hispanics to report their ethnicity as Hispanic and to report no race at all.

Race coding categories are not consistent over time. From 1948 through 1998 only one race was reported for each respondent, and from 2000 through 2012 multiple mentions were allowed. In 1948-1964 the race category of “Other” includes Hispanic as well as other identities such as Asian, so it is not possible to reliably distinguish race and ethnicity in these years. In 1966 and later years race and ethnicity can be distinguished, although the only Hispanic categories in 1966, 1968, and 1970 were Puerto Rican and Mexican (or “Mexican-American, Chicano” in 1970). In 1972, 1974, and 1976 the ethnicity data come from an open-ended question that may not reliably distinguish Hispanic categories. From 1978 forward we have reasonably consistent and reliable distinctions among Mexican-American, Puerto Rican, and “other Hispanic” backgrounds. From 1988 forward, we can reliably identify Cubans as well, though Cubans may also have self-reported this identity in the open-ended question in use since 1972.

The concept of race being measured changes over time. The instrumentation used for the project’s first 50 years, through 1998, measures racial appearances as judged by the interviewer. This measurement method reflects a conception of race as an outwardly observable and obvious characteristic. We have no information about the reliability of these judgments or their correspondence to respondent’s racial self-identification. The newer questions ask the respondents to report their own racial identity. Moreover, the newer questions allow multiple identities.

The electronic data files can be misleading about the data collection. There is information in the original questionnaires that has not been recorded in the electronic data files. For example, the 1984 paper questionnaire allows the interviewer to indicate whether the respondent is “definitely” or “probably” white, black, American Indian or Alaskan Native, or Asian or Pacific Islander. But the 1984 dataset

records no distinction between definite and probable classifications. The Hispanic ethnicity coding presents options of 1) Yes, Mexican American, 2) Yes, Puerto Rican, 3) Yes, other Hispanic, 5) No. This numbered sequence implies that the response options were presented in this order, but they were not; the left-right orientation of the response options on the questionnaire presented “No” first but numbered it fifth.

Difficulties with the CDF race and ethnicity variables

The current CDF has no race data for 2008.

CDF race variable labels make no reference to Hispanic ethnicity so users cannot tell if “White” includes Hispanics or not.

VCF0105 (2-category race) combines missing data with substantive responses. The two valid codes do not describe the entire population. This means that if a user runs a weighted frequency on this variable with missing data excluded the result gives false population proportions for the two categories the variable identifies, the black and white segments of the population. The variable seems useless and misleading.

It appears that the 3-category race variable must make inconsistent definitions of white and black over time, given that the source data for white and black include Hispanics in some years and exclude them in others. In some years, “white” means “white non-Hispanic” and in other years it includes Hispanics. Same for black. Same for the 6-category race variable.

It is desirable to present race/ethnicity data in two ways. One is to distinguish categories of white non-Hispanic, black non-Hispanic, other non-Hispanic, and Hispanic. This presentation is consistent with a lot of demographic reporting. The other is to have dummy variables for white, black, Hispanic, Asian, and other identities. This presentation is flexible and allows data analysts to combine categories as they see fit. Unfortunately the later approach is not really viable on the CDF because data have only been collected in a manner appropriate to this sort of presentation in the last two Time Series surveys.

Difficulties with summarizing the data we have

There is no fully satisfactory way to reconcile the changes in measurement over time and summarize race/ethnicity. Five potential options are as follows:

Option 1: dummy variables: create variables indicating if the respondent is or is not white, is or is not black, is or is not Hispanic, etc.

This option is problematic because for the years 1948-1976 we do not know the race of any Hispanic respondents. Also, from 1948-1964, Hispanics are combined with “other” races. We could provide these variables for 1978-2012.

Option 2: 4-category race/ethnicity summary: create a variable with these categories:

- White non-Hispanic
- Black non-Hispanic
- Hispanic
- Other or multiple races, non-Hispanic

This option is problematic because for the years 1948-1964 we cannot differentiate Hispanic from “Other or multiple races.” We could provide this variable for 1966-2012, although the distributions of race-ethnicity across the categories are subject to substantial measurement error and will vary over time due to changes in instrumentation.

Option 3: race/ethnicity summary combining Hispanics with “other” : create a variable with these categories:

- White non-Hispanic
- Black non-Hispanic
- Other or Hispanic
- Missing

This option, which may constitute CDF variable VCF0106, is problematic because it is insufficiently detailed for many analysts’ purposes.

Option 4: 6-category race/ethnicity summary with these categories:

- White non-Hispanic
- Black non-Hispanic
- Asian or Pacific Islander non-Hispanic
- American Indian or Alaska Native non-Hispanic
- Hispanic
- Other or multiple races, non-Hispanic

This option, which may constitute CDF variable VCF0106a with more precise labels, is problematic because Asians, Hispanics, American Indians, and other groups cannot be differentiated in 1948-1964. It is therefore not accurate in those years, as respondents in these groups would be coded missing or dropped into the “Other” category.

Option 5: race/ethnicity summary with time-variant categories:

- White non-Hispanic (1948-2012)
- Black non-Hispanic (1948-2012)
- Asian or Pacific Islander, non-Hispanic (1966-2012)
- American Indian or Alaska Native non-Hispanic (1966-2012)
- Hispanic (1966-2012)
- Other or multiple races, non-Hispanic (1968-2012)
- Non-white and non-black (1948-1964)

This option incorporates what we know from the questionnaire over time. It is problematic because it is inelegant and complicated. However, it reflects the state of the data and it facilitates the recoding of categories according to analysts’ preferences.

Proposal

I propose the following for the next CDF release.

1. Create a Race/Ethnicity summary with the following categories and replace the current 6-category race/ethnicity summary with it:
 - White non-Hispanic (1948-2012)
 - Black non-Hispanic (1948-2012)
 - Asian or Pacific Islander, non-Hispanic (1966-2012)
 - American Indian or Alaska Native non-Hispanic (1966-2012)
 - Hispanic (1966-2012)
 - Other or multiple races, non-Hispanic (1968-2012)
 - Non-white and non-black (1948-1964)
2. Permanently drop 2-category race, VCF0105, from the CDF.
3. Re-label 3-category race, VCF0106, as follows (derived from the summary):
 - White non-Hispanic
 - Black non-Hispanic
 - Other
 - Missing
4. Create 4-category race/ethnicity summary with the following categories derived from the main summary :
 - White non-Hispanic
 - Black non-Hispanic
 - Hispanic
 - Other or multiple races, non-Hispanic
 - Missing, pre-1966 data
 - Missing, DK/REF/NA
5. Update the Hispanic ethnicity variables, VCF0107, VCF0108, and VCF0109, to the extent possible with the latest data.

Coding Details for Proposed Summaries

Create the Race/Ethnicity summary with codes as follows.

1. White non-Hispanic (1948-2012)
2. Black non-Hispanic (1948-2012)
3. Asian or Pacific Islander, non-Hispanic (1966-2012)
4. American Indian or Alaska Native non-Hispanic (1966-2012)
5. Hispanic (1966-2012)
6. Other or multiple races, non-Hispanic (1968-2012)
7. Non-white and non-black (1948-1964)

For 1948-1964

Use the interviewer observation of the race of the respondent.

If observed race = "White" then race/ethnicity = 1 (white non-Hispanic).

If observed race = "Negro" then race/ethnicity = 2 (black non-Hispanic).

If observed race = "Other" then race/ethnicity = 7 (non-white and non-black).

Race/ethnicity summary is missing if there is item nonresponse on the interviewer observation.

For 1966-1970

Use the interviewer observation of the race of the respondent.

If "White" race/ethnicity = 1.

If "Negro" or "Black" race/ethnicity = 2.

If "Oriental" race/ethnicity = 3.

If "American Indian" race/ethnicity = 4.

If "Puerto Rican" race/ethnicity = 5.

If "Mexican-American, Chicano" or "Mexican" race/ethnicity = 5.

If "Other" race/ethnicity = 6.

Race/ethnicity summary is missing if there is item nonresponse on the interviewer observation.

For 1972-1976

Use the interviewer observation of the race of the respondent and for Hispanic origin also use the respondent's answer to "In addition to being American, what do you consider your main ethnic group or nationality group?" First code based on observed race.

If "White" race/ethnicity = 1.

If "Negro" or "Black" race/ethnicity = 2.

If "Oriental" race/ethnicity = 3.

If "American Indian" race/ethnicity = 4.

If "Puerto Rican" race/ethnicity = 5.

If "Mexican-American, Chicano" or "Mexican" race/ethnicity = 5.

If "Other" race/ethnicity = 6.

Next, if reported ethnic group or nationality group includes any Hispanic origin, including Spanish, Mexican, Mexican-American, Chicano, Puerto Rican, etc., race/ethnicity = 5, regardless of race.

Race/ethnicity summary is missing if there is item nonresponse on the interviewer observation.

For 1978-1986

Use the interviewer observation of the race of the respondent and for Hispanic origin also use the interviewer's observation of Hispanic origin and also use the respondent's answer to "In addition to being American, what do you consider your main ethnic group or nationality group?" First code based on observed race.

If "White" race/ethnicity = 1.

If "Black" race/ethnicity = 2.

If "Asian or Pacific Islander" race/ethnicity = 3.

If "American Indian or Alaska Native" race/ethnicity = 4.

If “Other” race/ethnicity = 6.

Next, if the interviewer observation of Hispanic origin is “Yes” for any Hispanic origin, race/ethnicity = 5.

Last, if reported ethnic group or nationality group includes any Hispanic origin, including Spanish, Mexican, Mexican-American, Chicano, Puerto Rican, etc., race/ethnicity = 5, regardless of race and regardless of the interviewer observation.

Race/ethnicity summary is missing if 1) there is item nonresponse on the interviewer observation of race together with no indication that R is Hispanic (non-response or non-Hispanic for both the interviewer observation whether R is Hispanic and the main ethnic group question) or if 2) race is not item nonresponse but both indicators for Hispanic status are nonresponse.

For 1988-1998

Use the interviewer observation of the race of the respondent and for Hispanic origin use the respondent’s answer to “In addition to being American, what do you consider your main ethnic group or nationality group?” and also use the respondent’s answer to “Are you of Spanish or Hispanic origin or descent? First code based on observed race.

If “White” race/ethnicity = 1.

If “Black” race/ethnicity = 2.

If “Asian or Pacific Islander” race/ethnicity = 3.

If “American Indian or Alaska Native” race/ethnicity = 4.

If “Other” race/ethnicity = 6.

Next, if reported ethnic group or nationality group includes any Hispanic origin, including Spanish, Mexican, Mexican-American, Chicano, Puerto Rican, etc., race/ethnicity = 5, regardless of race.

Last, if the answer to “Are you of Spanish or Hispanic origin or descent?” is “Yes,” race/ethnicity = 5, regardless of race.

Race/ethnicity summary is missing if 1) there is item nonresponse on the interviewer observation of race together with no indication that R is Hispanic (non-response or non-Hispanic for both the main ethnic group and Hispanic descent questions) or if 2) race is not item nonresponse but both indicators for Hispanic status are nonresponse.

For 2000-2008

Use the responses to these questions: “What racial or ethnic group or groups best describes you?” “In addition to being American, what do you consider your main ethnic group or nationality group?” “Are you of Spanish or Hispanic origin or descent?”

First code based on reported racial or ethnic groups:

If “Hispanic” or “Hispanic or Latino” was mentioned, race/ethnicity = 5;

Else if “Hispanic” and “Hispanic or Latino” were not mentioned, and more than one group was mentioned, race/ethnicity = 6;

Else if “Hispanic” and “Hispanic or Latino” were not mentioned, and only one group was mentioned, code as follows:

If group = “White” race/ethnicity = 1.

If group = “Black” race/ethnicity = 2.

If group = “Asian” race/ethnicity = 3.

If group = “Native American” race/ethnicity = 4.

If group = “Other” race/ethnicity = 6.

Next, if reported ethnic group or nationality group includes any Hispanic origin, including Spanish, Mexican, Mexican-American, Chicano, Puerto Rican, etc., race/ethnicity = 5, regardless of race.

Last, if the answer to “Are you of Spanish or Hispanic origin or descent?” is “Yes,” race/ethnicity = 5, regardless of race.

Race/ethnicity summary is missing if 1) there is item nonresponse for R’s race together with no indication that R is Hispanic (non-response or non-Hispanic for both the main ethnic group and Hispanic descent questions) or if 2) race is neither nonresponse nor ‘Hispanic or Latino’ and both remaining indicators for Hispanic status (main ethnic group and Hispanic descent questions) are nonresponse.

For 2012 and continuing years

Use reported race categories (“I am going to read you a list of five race categories. Please choose one or more races that you consider yourself to be.”) and reported Spanish, Hispanic, or Latino identity. (Do not use ethnic or nationality group items or other items that would make the measurement incomparable to Census Bureau methods.)

If respondent answered “Yes” to “Are you Spanish, Hispanic, or Latino?” race/ethnicity = 5.

Race/ethnicity is also coded to 5 if R provides an 'Other-specify' response to the race question which is an Hispanic mention (special code: "Other response given as an Hispanic mention").

If respondent race/ethnicity is not coded 5, race/ethnicity is coded based on race report as follows:

If more than one race was chosen, race/ethnicity = 6.

If only response was “White” race/ethnicity = 1.

If only response was “Black or African-American” race/ethnicity = 2.

If only response was “Asian” race/ethnicity = 3.

If only response was “Native Hawaiian or other Pacific Islander” race/ethnicity = 3.

If only response was “American Indian or Alaska Native” race/ethnicity = 4.

If only response was “Other” race/ethnicity = 6.

Race/ethnicity summary is missing if 1) there is item nonresponse for R's race together with response 'No' or item nonresponse for the question asking if R is Spanish, Hispanic, or Latino, or if 2) race is not item nonresponse and does not include an Other-specify Hispanic mention but the question asking if R is Spanish, Hispanic, or Latino is nonresponse.

Create the three-category race summary, VCF0106, by recoding the main race/ethnicity summary as follows.

If race/ethnicity = 1 VCF0106 = 1 (white non-Hispanic)

If race/ethnicity = 2 VCF0106 = 2 (black non-Hispanic)

If race/ethnicity = 3 VCF0106 = 3 (other)

If race/ethnicity = 4 VCF0106 = 3 (other)

If race/ethnicity = 5 VCF0106 = 3 (other)

If race/ethnicity = 6 VCF0106 = 3 (other)

If race/ethnicity = 7 VCF0106 = 3 (other)

If race/ethnicity = missing VCF0106 = missing

Create the four-category race/ethnicity summary by recoding the main race/ethnicity summary as follows:

If race/ethnicity = 1 4_category_summary = white non-Hispanic

If race/ethnicity = 2 4_category_summary = black non-Hispanic

If race/ethnicity = 3 4_category_summary = other or multiple races, non-Hispanic

If race/ethnicity = 4 4_category_summary = other or multiple races, non-Hispanic

If race/ethnicity = 5 4_category_summary = Hispanic

If race/ethnicity = 6 4_category_summary = other or multiple races, non-Hispanic

If race/ethnicity = 7 4_category_summary = Missing, pre-1966 data

If race/ethnicity = missing 4_category_summary = Missing, DK/REF/NA

Appendix1: Codebook entry for CDF race/ethnicity variables.

```
=====
VCF0105      Respondent Race 2-category
=====
```

QUESTION:

1948-1998: INTERVIEWER OBSERVATION OF RACE
2000 AND LATER: What racial or ethnic group or groups best describes you?
1972 AND LATER (exc. 2002): In addition to being American, what do you
consider your main ethnic group or nationality group?

VALID_CODES:

1. White
2. Black

MISSING_CODES:

0. NA; "other"; no Pre IW; short-form 'new' Cross
Section or Spanish language (1992)
Inap. question not used

NOTES:

Collapsed from VCF0106a.

WEIGHT:

VCF0009/VCF0009a

TYPE:

Numeric Dec 0

SOURCE_VARS:

1948: 480046
1952: 520129
1954: 540027
1956: 560172
1958: 580171
1960: 600119
1962: 620097
1964: 640183
1966: 660237
1968: 680264
1970: 700391(type 0)
1972: 720406,720425
1974: 742528,742554
1976: 763489,763513
1978: 780608,780631,780632
1980: 800695,800721
1982: 820736,820763
1984: 840689,840708
1986: 860739,860756
1988: 880412,880534
1990: 900549,900670
1992: 924116,924202
1994: 941412,941435
1996: 960067,960703
1998: 980654,980673
2000: 001006a,001008
2002: 023150
2004: 043299a,043301a

=====
VCF0106

Respondent Race 3-category
=====

QUESTION:

1948-1998: INTERVIEWER OBSERVATION OF RACE

2000 AND LATER: What racial or ethnic group or groups best describes you?

1972 AND LATER (exc. 2002): In addition to being American, what do you
consider your main ethnic group or nationality group?

VALID_CODES:

1. White
2. Black
3. Other

MISSING_CODES:

9. DK; NA; no Pre IW; short-form 'new' Cross Section
or Spanish language (1992)
Inap. question not used

NOTES:

Collapsed from VCF0106a.

WEIGHT:

VCF0009/VCF0009a

TYPE:

Numeric Dec 0

SOURCE_VARS:

1948: 480046
1952: 520129
1954: 540027
1956: 560172
1958: 580171
1960: 600119
1962: 620097
1964: 640183
1966: 660237
1968: 680264
1970: 700391(type 0)
1972: 720406,720425
1974: 742528,742554
1976: 763489,763513
1978: 780608,780631,780632
1980: 800695,800721
1982: 820736,820763
1984: 840689,840708
1986: 860739,860756
1988: 880412,880534
1990: 900549,900670
1992: 924116,924202
1994: 941412,941435
1996: 960067,960703
1998: 980654,980673
2000: 001006a,001008
2002: 023150
2004: 043299a,043301a

=====
VCF0106a Respondent Race 6-category
=====

QUESTION:

1948-1998: INTERVIEWER OBSERVATION OF RACE
2000 AND LATER: What racial or ethnic group or groups best describes you?
1972 AND LATER (exc. 2002): In addition to being American, what do you
consider your main ethnic group or nationality group?

VALID_CODES:

- 1. White
2. Black
3. Asian
4. Native American
5. Hispanic
7. Other

MISSING_CODES:

- 0. DK; NA; no Pre IW; short-form 'new' Cross Section
 or Spanish language (1992)
Inap. question not used

NOTES:

GENERAL NOTE:

When variables for IWR observation of race were present in both the Pre and Post, the Pre variable has been utilized (1960,1972,1976,1980-1986). In years prior to 1972, race was identified by interviewer observation alone. Beginning in 1972, respondents were also asked for ethnic self-identification. VCF0106a includes only the IWR observation race data for 1948-1970, but reclassifies IWR observation race of 1972-1998 respondents who were observed as White or Other if their primary (first) mention for ethnic self-identification was a Black, Asian, Native American, or Hispanic mention. (These reclassifications of race from specific mentions for ethnic/national self-identification are in conformity with available Census definitions).

1960 NOTE:

The Pre variable incorporated data from Post race variable V600249 to establish the race of 22 no-pre cases.

1992 NOTE:

Interviewer observation of race was not part of the 'short-form' questionnaire; for short-form panel cases, however, 1990 study data were incorporated.

2000 NOTE:

Interviewer observation was not utilized for identification of respondent race. Instead, respondents were first asked to self-identify their own racial group and were then asked for their ethnic self-identification. Respondents with a primary (first) mention for race which was not Black, Asian, Native American or Hispanic had race reclassified in VCF0106a for race from primary ethnic self-identification in the same manner as for 1972-1998.

2002 NOTE:

The ethnicity question was not asked and this variable was coded from the racial self-identification mentions. Since order of mentions was not determined, combined mention of white and other racial group has been recoded to the nonwhite racial group.

2004 NOTE:

same as 2000.

1 case of middle-eastern respondent who identified self as Asian (in British fashion) has been coded to 1.

WEIGHT:

VCF0009/VCF0009a

TYPE:

Numeric Dec 0

SOURCE_VARS:

1948: 480046
1952: 520129
1954: 540027
1956: 560172
1958: 580171
1960: 600119
1962: 620097
1964: 640183
1966: 660237
1968: 680264
1970: 700391(type 0)
1972: 720406,720425
1974: 742528,742554
1976: 763489,763513
1978: 780608,780631,780632
1980: 800695,800721
1982: 820736,820763
1984: 840689,840708
1986: 860739,860756
1988: 880412,880534
1990: 900549,900670
1992: 924116,924202
1994: 941412,941435
1996: 960067,960703
1998: 980654,980673
2000: 001006a,001008
2002: 023150
2004: 043299a,043301a

=====
VCF0107 R Hispanic Origin Type
=====

QUESTION:

1988 AND LATER: In addition to being American, what do you consider your main ethnic group or nationality group? (IF NO HISPANIC GROUP MENTIONED:) Are you of Spanish or Hispanic origin or descent? IF YES:) Please look at the booklet and tell me which category best describes your Hispanic origin.

VALID_CODES:

1. Yes, Mexican-American, Chicano
2. Yes, Puerto Rican
3. Yes, other Hispanic
4. Yes, Hispanic but DK/NA type
7. No, not Hispanic

MISSING_CODES:

8. DK if Hispanic
9. NA if Hispanic
0. Short-form 'new' Cross Section (1992)
Inap. question not used

NOTES:

SEE ALSO VCF0108.

GENERAL NOTE:

RESPONDENT BOOKLET SHOWN TO R (exc. 2000 telephone and 2002:) .
Beginning in 1988, the respondent was asked within the series of
ethnicity questions [SEE VCF0109] whether s/he is Hispanic. Prior to
1988, information about Hispanic origin was determined only by
interviewer observation.

1978 NOTE:

A large number (171) of NAs appeared in the Study data.

1980 NOTE:

A Hispanic origin variable was present in both the pre and post. The
pre version is represented here.

1984 NOTE:

same as 1980

1992 NOTE:

This was not part of the 'short-form' questionnaire; for short-form
panel cases, however, 1990 study data were incorporated.

1994 NOTE:

For panel cases, 1992 data were incorporated.

2002 NOTE:

For panel cases, 2000 data were incorporated.

WEIGHT:

VCF0009/VCF0009a

TYPE:

Numeric Dec 0

SOURCE_VARS:

1978: 780632
1980: 800722
1982: 820764
1984: 840709
1986: 860757
1988: 880540,880541
1990: 900676,900677
1992: 924122,924123
1994: 941418,941419
1996: 960708,960709
1998: 980659,980660
2000: 001012,001013
2002: 023152
2004: 043306
2008: 083256a

=====
VCF0108 R Hispanic Origin
=====

QUESTION:

1988 AND LATER: In addition to being American, what do you consider
your main ethnic group or nationality group? (IF NO HISPANIC GROUP
MENTIONED:) Are you of Spanish or Hispanic origin or descent?
IF YES:) Please look at the booklet and tell me which category best
describes your Hispanic origin.

VALID_CODES:

1. Yes, R is Hispanic
2. No, R is not Hispanic

MISSING_CODES:

8. DK

9. NA
0. Short-form 'new' Cross Section (1992)
Inap. question not used

NOTES:

Collapsed from VCF0107.

WEIGHT:

VCF0009/VCF0009a

TYPE:

Numeric Dec 0

SOURCE_VARS:

1978: 780632
1980: 800722
1982: 820764
1984: 840709
1986: 860757
1988: 880540,880541
1990: 900676,900677
1992: 924122,924123
1994: 941418,941419
1996: 960708,960709
1998: 980659,980660
2000: 001012,001013
2002: 023151
2004: 043305,043306
2008: 083256a

=====
VCF0109 R Ethnicity
=====

QUESTION:

In addition to being an American what do you consider your main ethnic group or nationality group?
(IF MORE THAN 1 GROUP MENTIONED:) With which of these groups do you most closely identify?

VALID_CODES:

010-874. Ethnicity as coded

MISSING_CODES:

877. 'American'; 'Just American'; none; neither
(response to 'choice' question); NA
888. DK
000. Short-form 'new' Cross Section (1992)
Inap. question not used

NOTES:

See ethnicity codes- Appendix "ETHNICITY"

GENERAL NOTE:

This represents the single mention of respondents who indicated only 1 ethnic group and the 'most close' group (followup question) of respondents who indicated more than 1 ethnic group. There is no code to single out Rs who volunteered to assert their 'American' identity. This is because in most years these Rs were combined with NA and/or 'none' (no ethnic identity).

1978 NOTE:

A version of this question which included a (filter) lead-in was used:
'In addition to being an American, is there another nationality group
or ethnic group that you feel you belong to? [IF YES] What...'
Because of this format, the 1978 distributions varied dramatically from
other years, and 1978 data have not been included.

1992 NOTE:

This was not part of the 'short-form' questionnaire; for short-form
panel cases, however, 1990 study data were incorporated.

1994 NOTE:

For panel cases, 1992 data were incorporated.

1996 NOTE:

In a few cases R identified 2 or more nationality/ethnicity groups but
the interviewer did not ask the 'choice' followup (due to incorrect
completion of the IWR checkpoint); since R's chosen nationality/
ethnicity remained undetermined, these cases (30) were considered
missing data (code 99 here).

2004 NOTE:

For respondents who gave multiple mentions but did not choose one
ethnic group as closest, 1st mention is coded.

2008 NOTE:

For respondents who gave multiple mentions but did not choose one
ethnic group as closest, 1st mention is coded.

WEIGHT:

VCF0009/VCF0009a

TYPE:

Numeric Dec 0

SOURCE_VARS:

1972: 720406
1974: 742528
1976: 763489
1980: 800695,800696,800697
1982: 820739
1984: 840692
1986: 860742
1988: 880537
1990: 900673
1992: 924116,924119
1994: 941415
1996: 960706
1998: 980657
2000: 001008
2004: 043303x
2008: 083252a,083253,083254

MemoAppendix_race-ethrcode.txt

```
data library.cdf_raceeth ;
  set library.all_tsrace8 ;
* Input file has all relevant race vars 1948-2012 merged into a single file ;
* ===== ;
* CREATE MAIN CDF SUMMARY VCF0105a ;
* Code written for maximum transparency ;
* ===== ;
* ;
* ----- ;
* Years 1948-1964 have only 1 variable (IWR observation of race) with ;
* 3 codes: White (1), Black (2), Other (3,5, or 7). ;
* These are the only years from which cases are coded to 7 ("Nonwhite ;
* and nonblack") for VCF0105a. ;
* Cases missing in VCF0105a are missing in the year-level ;
* variable. ;
* ;
if year = 1948 then do ;
  if V480046 in(1,2 ) then VCF0105a = V480046 ;
  if V480046 = 3 then VCF0105a = 7 ;
  go to MD ;
end ;
if year = 1952 then do ;
  if V520129 in(1,2,7) then VCF0105a = V520129 ;
end ;
if year = 1956 then do ;
  if V560172 in(1,2,7) then VCF0105a = V560172 ;
  go to MD ;
end ;
if year = 1958 then do ;
  if V580171 in(1,2,7) then VCF0105a = V580171 ;
  go to MD ;
end ;
if year = 1960 then do ;
  if V600119 in(1,2,7) then VCF0105a = V600119 ;
  go to MD ;
end ;
if year = 1962 then do ;
  if V620097 in(1,2) then VCF0105a = V620097 ;
  if V620097 = 5 then VCF0105a = 7 ;
  go to MD ;
end ;
if year = 1964 then do ;
  if V640183 in(1,2,7) then VCF0105a = V640183 ;
  go to MD ;
end ;
* ;
* ----- ;
* Years 1966-1970 have only 1 variable (IWR observation of race) with ;
* codes: White (1), Black (2), Puerto Rican (3), Mexican (4), Oriental (5), ;
* American Indian (6), Other (7). ;
* Cases missing in VCF0105a are missing in the year-level ;
* variable. ;
* ;
if year = 1966 then do ;
```

MemoAppendix_race-ethrcode.txt

```

if V660237 in(1,2) then VCF0105a = V660237 ;
if V660237 in(3,4) then VCF0105a = 5 ;
if V660237 = 5 then VCF0105a = 3 ;
if V660237 = 6 then VCF0105a = 4 ;
if V660237 = 7 then VCF0105a = 6 ;
go to MD ;
end ;
if year = 1968 then do ;
if V680264 in(1,2) then VCF0105a = V680264 ;
if V680264 in(3,4) then VCF0105a = 5 ;
if V680264 = 5 then VCF0105a = 3 ;
if V680264 = 6 then VCF0105a = 4 ;
if V680264 = 7 then VCF0105a = 6 ;
go to MD ;
end ;
if year = 1970 then do ;
if V700391 in(1,2) then VCF0105a = V700391 ;
if V700391 in(3,4) then VCF0105a = 5 ;
if V700391 = 5 then VCF0105a = 3 ;
if V700391 = 6 then VCF0105a = 4 ;
if V700391 = 7 then VCF0105a = 6 ;
go to MD ;
end ;
* ;
* ----- ;
* Years 1972-1976 have a variable (IWR observation of race) with the same ;
* codes as 1966-1970. ;
* In addition, R was asked for main ethnicity or nationality ("in ;
* addition to American"). 1 mention for ethnicity-nationality was coded ;
* using the State-Country master code. ;
* Ethnicity-nationality mentions coded to Hispanic for VCF0105a ;
* include: 182 Puerto Rico, 219 Mexico, 259 South America, 353 Spain. ;
* South America is included in correspondence with Census Bureau ;
* categorization. (Only 2 cases are 259 South America for each ;
* year 1972-1976.) ;
* Cases missing in the CDF summary VCF0105a are those having missing ;
* data for the race variable. Because the race variable includes ;
* Hispanic codes, missing data for the ethnicity-nationality mention ;
* does not by itself result in missing data in VCF0105a. ;
if year = 1972 then do ;
if V720425 in(1,2) then VCF0105a = V720425 ;
if V720425 in(3,4) then VCF0105a = 5 ;
if V720425 = 5 then VCF0105a = 3 ;
if V720425 = 6 then VCF0105a = 4 ;
if V720425 = 7 then VCF0105a = 6 ;
if V720406 in(182,219,259,353) then VCF0105a = 5 ;
go to MD ;
end ;
if year = 1974 then do ;
if V742554 in(1,2) then VCF0105a = V742554 ;
if V742554 in(3,4) then VCF0105a = 5 ;
if V742554 = 5 then VCF0105a = 3 ;
if V742554 = 6 then VCF0105a = 4 ;
if V742554 = 7 then VCF0105a = 6 ;

```

MemoAppendix_race-ethrcode.txt

```

if V742528 in(182,219,259,353) then VCF0105a = 5 ;
go to MD ;
end ;
if year = 1976 then do ;
  if V763513 in(1,2)      then VCF0105a = V763513 ;
  if V763513 in(3,4)    then VCF0105a = 5 ;
  if V763513 = 5        then VCF0105a = 3 ;
  if V763513 = 6        then VCF0105a = 4 ;
  if V763513 = 7        then VCF0105a = 6 ;
  if V763489 in(182,219,259,353) then VCF0105a = 5 ;
go to MD ;
end ;
* ;
* ----- ;
* Years 1978-1986 have 4 variables, including IWR observation of race ;
* with codes White (1), Black (2), American Indian or Alaskan Native (3), ;
* Asian or Pacific Islander (4), Other (7). (The 'race' variable no longer ;
* includes any Hispanic-specific categories as in 1966-1976.) ;
* The open-end ethnicity-nationality question also continues to be asked, ;
* however a new 'Nationality' master code is applied and 2 mentions are ;
* coded (the Nationality master code is closely parallel to, but not ;
* identical to, the State-Country master code used 1972-1976).
* Ethnicity-nationality mentions coded to Hispanic in VCF0105a include: ;
* 04 Mexican, 05 Central American, 08 Cuban, 09 Dominican Republic, ;
* 12 Puerto Rican, 16 South American, 62 Spanish, 88 Chicano or Mexican- ;
* American.
* The fourth variable is an IWR observation whether or not R is Hispanic, ;
* coded Yes Mexican-American (1), Yes Puerto Rican (2), Yes Other Hispanic ;
* (3) and No (5). ;
* Cases missing in the CDF summary VCF0105a are those having: ;
* 1) missing data for race together with no Hispanic status indicated (all ;
* variables used to indicate Hispanic status are missing or non-Hispanic), or ;
* 2) race not missing data but missing data for all variables used to ;
* identify Hispanic status (a Hispanic respondent is classified to ;
* Hispanic in VCF0105a even when race is missing). ;
* ;
* 1978 NOTE: 1978 is the only year in which the question about main ;
* ethnicity-nationality is preceded by a filter question (V780607) asking ;
* whether or not R identifies with an ethnic or nationality group (Yes/No). ;
* The answer No to this filter question corresponds to ethnicity-nationality ;
* mention 'None' in subsequent years, rather than to missing data for ;
* ethnicity-nationality. ;
if year = 1978 then do ;
  if V780631 in(1,2)      then VCF0105a = V780631 ;
  if V780631 = 3          then VCF0105a = 4 ;
  if V780631 = 4          then VCF0105a = 3 ;
  if V780631 = 7          then VCF0105a = 6 ;
  if V780632 in(1,2,3)    then VCF0105a = 5 ;
  if V780608 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
  if V780609 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
  if V780631 = 9 and V780632 not in(1,2,3)
    and V780608 not in(04,05,08,09,12,16,62,88)
    and V780609 not in(04,05,08,09,12,16,62,88) then VCF0105a = 9 ;

```

MemoAppendix_race-ethrcode.txt

```
* Race not missing but Hispanic status items missing data ;
if V780631 ne 9 and V780632 in(8,9)
    and (V780607 in(8,9) or V780608 in(98,99) ) then VCF0105a = 9 ;
go to SKIP ;
end ;
if year = 1980 then do ;
    if V800721 in(1,2)          then VCF0105a = V800721 ;
    if V800721 = 3              then VCF0105a = 4 ;
    if V800721 = 4              then VCF0105a = 3 ;
    if V800721 = 7              then VCF0105a = 6 ;
    if V800722 in(1,2,3)       then VCF0105a = 5 ;
    if V800695 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V800696 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V800721 = 9 and V800722 not in(1,2,3)
    and V800695 not in(04,05,08,09,12,16,62,88)
    and V800696 not in(04,05,08,09,12,16,62,88) then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
if V800721 ne 9 and V800722 in(8,9)
    and V800695 in(98,99)          then VCF0105a = 9 ;
go to SKIP ;
end ;
if year = 1982 then do ;
    if V820763 in(1,2)          then VCF0105a = V820763 ;
    if V820763 = 3              then VCF0105a = 4 ;
    if V820763 = 4              then VCF0105a = 3 ;
    if V820763 = 7              then VCF0105a = 6 ;
    if V820764 in(1,2,3)       then VCF0105a = 5 ;
    if V820736 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V820737 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V820763 = 9 and V820764 not in(1,2,3)
    and V820736 not in(04,05,08,09,12,16,62,88)
    and V820737 not in(04,05,08,09,12,16,62,88) then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
if V820763 ne 9 and V820764 in(8,9)
    and V820736 in(98,99)          then VCF0105a = 9 ;
go to SKIP ;
end ;
if year = 1984 then do ;
    if V840708 in(1,2)          then VCF0105a = V840708 ;
    if V840708 = 3              then VCF0105a = 4 ;
    if V840708 = 4              then VCF0105a = 3 ;
    if V840708 = 7              then VCF0105a = 6 ;
    if V840709 in(1,2,3)       then VCF0105a = 5 ;
    if V840689 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V840690 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V840708 = 9 and V840709 not in(1,2,3)
    and V840689 not in(04,05,08,09,12,16,62,88)
    and V840690 not in(04,05,08,09,12,16,62,88) then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
if V840708 ne 9 and V840709 in(8,9)
    and V840689 in(98,99)          then VCF0105a = 9 ;
```

MemoAppendix_race-ethrcode.txt

```

go to SKIP ;
end ;
if year = 1986 then do ;
  if V860756 in(1,2)      then VCF0105a = V860756 ;
  if V860756 = 3         then VCF0105a = 4 ;
  if V860756 = 4         then VCF0105a = 3 ;
  if V860756 = 7         then VCF0105a = 6 ;
  if V860757 in(1,2,3)   then VCF0105a = 5 ;
  if V860739 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
  if V860740 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
  if V860756 = 9 and V860757 not in(1,2,3)
    and V860739 not in(04,05,08,09,12,16,62,88)
    and V860740 not in(04,05,08,09,12,16,62,88) then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
  if V860756 ne 9 and V860757 in(8,9)
    and V860739 in(98,99)      then VCF0105a = 9 ;
go to SKIP ;
end ;
* ;
* ----- ;
* Years 1988-1998 have 4 variables, including 3 variables for the same ;
* IWR observation of race and open-end ethnicity-nationality mentions ;
* used in 1978-1986. However, IWR observation whether or not R is Hispanic;
* is no longer used as in 1978-1986. ;
* Instead, if R does not mention a Hispanic group for ethnicity- ;
* nationality, R is asked if he/she is of Hispanic origin or descent: ;
* Rs answering Yes (1) are coded to Hispanic in the CDF summary. ;
* Cases missing in the CDF summary VCF0105a are those having: ;
* 1) missing data for race together with no Hispanic status indicated ;
* (missing or non-Hispanic in all variables used to indicate Hispanic ;
* status), or ;
* 2) race not missing data but missing data for all variables used to ;
* identify Hispanic status (a Hispanic respondent is classified to ;
* Hispanic in VCF0105a even when race is missing). ;
* ;
if year = 1988 then do ;
  if V880412 in(1,2)      then VCF0105a = V880412 ;
  if V880412 = 3         then VCF0105a = 4 ;
  if V880412 = 4         then VCF0105a = 3 ;
  if V880412 = 7         then VCF0105a = 6 ;
  if V880534 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
  if V880535 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
  if V880540 = 1         then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
  if V880412 = 9 and V880534 not in(04,05,08,09,12,16,62,88)
    and V880535 not in(04,05,08,09,12,16,62,88)
    and V880540 ne 1      then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
  if V880412 ne 9 and V880534 in(98,99)
    and V880540 in(8,9)  then VCF0105a = 9 ;
go to SKIP ;
end ;
if year = 1990 then do ;

```

MemoAppendix_race-ethrcode.txt

```

if V900549 in(1,2)      then VCF0105a = V900549 ;
if V900549 = 3         then VCF0105a = 4 ;
if V900549 = 4         then VCF0105a = 3 ;
if V900549 = 7         then VCF0105a = 6 ;
if V900670 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if V900671 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if V900676 = 1         then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V900549 = 9 and V900670 not in(04,05,08,09,12,16,62,88)
    and V900671 not in(04,05,08,09,12,16,62,88)
    and V900676 ne 1    then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
if V900549 ne 9 and V900670 in(98,99)
    and V900676 in(8,9) then VCF0105a = 9 ;
go to SKIP ;
end ;
* 1992 NOTES: ;
* 1) Rs administered Spanish-language interviews (6,7 in V923011) ;
* do not have an IWR observation of race made in 1992 (0 in ;
* V924202). When these are panel cases (6 in V923011), however, ;
* IWR observation of race from 1990 (V900549) is used for VCF0105a ;
* 2) English-language short-form interviews (3,4 in V923011) do not ;
* include the open-end ethnicity-nationality question (0 in V924116). ;
* When these are panel cases (3 in V923011), however, ethnicity- ;
* nationality responses given in 1990 (V900670-V900671) are used for ;
* VCF0105a. ;
* 3) English-language short-form interviews (3,4 in V923011) do not ;
* include the question asking whether or not R is of Hispanic descent. ;
* When these are panel cases (3 in V923011), however, the 1990 ;
* data for this question are used for VCF0105a. (The 1992 ;
* variable for the question [V924122] includes the 1990 data for these ;
* cases.) ;
if year = 1992 then do ;
if V924202 in(1,2)      then VCF0105a = V924202 ;
if V924202 = 0         then VCF0105a = 9 ;
if V924202 = 3         then VCF0105a = 4 ;
if V924202 = 4         then VCF0105a = 3 ;
if V924202 = 7         then VCF0105a = 6 ;
if V923011 = 6 then do ;
if yr92_V900549 in(1,2) then VCF0105a = yr92_V900549 ;
if yr92_V900549 = 3    then VCF0105a = 4 ;
if yr92_V900549 = 4    then VCF0105a = 3 ;
if yr92_V900549 = 7    then VCF0105a = 6 ;
end ;
if V924116 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if V924117 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if V923011 = 3 then do ;
if yr92_V900670 in(04,05,08,09,12,16,62,88) or
    yr92_V900671 in(04,05,08,09,12,16,62,88)
    then VCF0105a = 5 ;
end ;
if V924122 = 1         then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V923011 not in(3,6) and V924202 eq 9

```

MemoAppendix_race-ethrcode.txt

```

and V924116 not in(04,05,08,09,12,16,62,88)
and V924117 not in(04,05,08,09,12,16,62,88)
and V924122 ne 1 then VCF0105a = 9 ;
if V923011 eq 6 and yr92_V900549 eq 9
and V924116 not in(04,05,08,09,12,16,62,88)
and V924117 not in(04,05,08,09,12,16,62,88)
and V924122 ne 1 then VCF0105a = 9 ;
if V923011 eq 3 and V924202 eq 9
and yr92_V900670 not in(04,05,08,09,12,16,62,88)
and yr92_V900671 not in(04,05,08,09,12,16,62,88)
and V924122 ne 1 then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
if V923011 not in(3,6) and V924202 ne 9
and V924116 in(98,99)
and V924122 in(8,9) then VCF0105a = 9 ;
if V923011 eq 6 and yr92_V900549 ne 9
and V924116 in(98,99)
and V924122 in(8,9) then VCF0105a = 9 ;
if V923011 eq 3 and V924202 ne 9
and yr92_V900670 in(98,99)
and V924122 in(8,9) then VCF0105a = 9 ;
go to SKIP ;
end ;
* 1994 NOTE: 1994 Panel cases (cases that were fresh cross in 1992) have been ;
* administered shorter 1994 interviews that do not include the open-end ethnicity- ;
* nationality item nor the followup question asking if R is of Hispanic descent. ;
* Status as fresh cross-section case or panel case in 1994 is coded in V941411. ;
* For 1994 Panel cases (2 in 941411), ethnicity-nationality is obtained from ;
* 1992 variables V924116 and V924117, and answers to the question about Hispanic ;
* descent are obtained from 1992 var V924122. ;
if year = 1994 then do ;
if V941435 in(1,2) then VCF0105a = V941435 ;
if V941435 = 3 then VCF0105a = 4 ;
if V941435 = 4 then VCF0105a = 3 ;
if V941435 = 7 then VCF0105a = 6 ;
if V941412 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if V941413 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if V941418 = 1 then VCF0105a = 5 ;
if V941411 = 2 then do ;
if yr94_V924116 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if yr94_V924117 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
if yr94_V924122 = 1 then VCF0105a = 5 ;
end ;
* Race missing and no Hispanic indication ;
if V941411 = 1 and V941435 = 9
and V941412 not in(04,05,08,09,12,16,62,88)
and V941413 not in(04,05,08,09,12,16,62,88)
and V941418 ne 1 then VCF0105a = 9 ;
if V941411 = 2 and V941435 = 9
and yr94_V924116 not in(04,05,08,09,12,16,62,88)
and yr94_V924117 not in(04,05,08,09,12,16,62,88)
and yr94_V924122 ne 1 then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
if V941411 = 1 and V941435 ne 9

```


MemoAppendix_race-ethrcode.txt

```

    and V941412 in(98,99)
    and V941418 in(8,9)      then VCF0105a = 9 ;
if V941411 = 2 and V941435 ne 9
    and yr94_V924116 in(98,99)
    and yr94_V924122 in(8,9)  then VCF0105a = 9 ;
go to SKIP ;
end ;
* 1996 NOTE: 1996 Panel cases (cases previously interviewed in 1994 ;
* some of whom were also interviewed in 1992) do not have IWR observation ;
* of race administered in 1996. However the 1996 release variable for ;
* IWR observation of race (V960067) includes, for 1996 panel cases, ;
* data from the preceding Time Series Study (1992 or 1994) in which ;
* IWR observation of race was most recently administered. ;
if year = 1996 then do ;
    if V960067 in(1,2)          then VCF0105a = V960067 ;
    if V960067 = 3              then VCF0105a = 4 ;
    if V960067 = 4              then VCF0105a = 3 ;
    if V960067 = 7              then VCF0105a = 6 ;
    if V960703 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V960704 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V960708 = 1              then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
    if V960067 = 9 and V960703 not in(04,05,08,09,12,16,62,88)
        and V960704 not in(04,05,08,09,12,16,62,88)
        and V960708 ne 1      then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
    if V960067 ne 9 and V960703 in(98,99)
        and V960708 in(8,9)  then VCF0105a = 9 ;
go to SKIP ;
end ;
if year = 1998 then do ;
    if V980673 in(1,2)          then VCF0105a = V980673 ;
    if V980673 = 3              then VCF0105a = 4 ;
    if V980673 = 4              then VCF0105a = 3 ;
    if V980673 = 7              then VCF0105a = 6 ;
    if V980654 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V980655 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    if V980659 = 1              then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
    if V980673 = 9 and V980654 not in(04,05,08,09,12,16,62,88)
        and V980655 not in(04,05,08,09,12,16,62,88)
        and V980659 ne 1      then VCF0105a = 9 ;
* Race not missing but Hispanic status items missing data ;
    if V980673 ne 9 and V980654 in(98,99)
        and V980659 in(8,9)  then VCF0105a = 9 ;
go to SKIP ;
end ;
* ----- ;
* Years 2000-2008 no longer employ IWR observation of race. ;
* {NOTE: there is no 2006 Time Series study}. ;
* However, the same open-end ethnicity-nationality question of prior years is present ;
* in 2000, 2004, and 2008 (2 mentions coded), and the question about Hispanic ;
* descent used 1980-1996 continues in all years 2000,2002,2004,2008.
* To identify race, R is asked, 'What racial or ethnic group or groups best ;

```

MemoAppendix_race-ethrcode.txt

```

* describe you?' and the maximum recorded responses given by any respondent ;
* determines the number of mention-variables in the study release. Major group ;
* response categories are Black (10 or 01), Asian (20 or 02), Native American ;
* (30 or 03), Hispanic or Latino (40 or 04), and White (50 or 05). ;
* Cases missing in the CDF summary VCF0105a are those having: ;
* 1) missing data for race together with no Hispanic status indicated in ;
* any other variable usable to indicate Hispanic status (missing or
* non-Hispanic), or ;
* 2) race is neither nonresponse nor 'Hispanic or Latino' but there is ;
* missing data for all other variables usable to identify Hispanic status ;
* (a Hispanic respondent is classified to Hispanic in VCF0105a even when race ;
* is missing). ;
* ;
* 2000 NOTE: For the release of race variables for the 2000 study (3 coded mentions), ;
* texts of 'Other-Specify' answers have been reviewed and, if found to constitute ;
* a group mention corresponding to codes 10,20,30,40, or 50 as described above,
* have been 'upcoded' appropriately. For example, an "African-American" response ;
* for Other-specify belongs to category 'Black' rather than to category 'Other'. ;
* Exceptions are "East Indian," "Pacific Islander," and "Jewish," which have been ;
* assigned to special (added) codes 75, 76, and 80, respectively, in the 2000 study ;
* data. ;
if year = 2000 then do ;
  * only 1 answer ;
  if V001006b = 00 and V001006c = 00 then do ;
    if V001006a = 10      then VCF0105a = 2 ;
    if V001006a = 20      then VCF0105a = 3 ;
    if V001006a = 30      then VCF0105a = 4 ;
    if V001006a = 40      then VCF0105a = 5 ;
    if V001006a = 50      then VCF0105a = 1 ;
    if V001006a in(75,76) then VCF0105a = 3 ;
    if V001006a in(60,79,80,90) then VCF0105a = 6 ;
  end ;
  * more than 1 answer ;
  if not (V001006b = 00 and V001006c = 00) then do ;
    if V001006a = 40 or V001006b = 40 or V001006c = 40      then VCF0105a = 5 ;
    else do ;
      black = 0 ; asian = 0 ; nativeam = 0 ; white = 0 ;
      if V001006a = 10 or V001006b = 10 or V001006c = 10 then black = 1 ;
      if V001006a = 20 or V001006b = 20 or V001006c = 20 then asian = 1 ;
      if V001006a = 75 or V001006b = 75 or V001006c = 75 then asian = 1 ;
      if V001006a = 76 or V001006b = 76 or V001006c = 76 then asian = 1 ;
      if V001006a = 30 or V001006b = 30 or V001006c = 30 then nativeam = 1 ;
      if V001006a = 50 or V001006b = 50 or V001006c = 50 then white = 1 ;
      numeth = black + asian + nativeam + white ;
      if numeth gt 1      then VCF0105a = 6 ;
      if numeth eq 1 and black = 1      then VCF0105a = 2 ;
      if numeth eq 1 and asian = 1      then VCF0105a = 3 ;
      if numeth eq 1 and nativeam = 1   then VCF0105a = 4 ;
      if numeth eq 1 and white = 1      then VCF0105a = 1 ;
      if numeth eq 0 and (V001006a in(60,80,90) or V001006b = 90 or V001006c = 90) then VCF0105a = 6 ;
    end ;
  end ;
  if V001008 in(04,05,08,09,12,16,62,88)      then VCF0105a = 5 ;
  if V001009 in(04,05,08,09,12,16,62,88)      then VCF0105a = 5 ;

```

MemoAppendix_race-ethrcode.txt

```

if V001012 = 1
    then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V001006a in(00,98,99) and V001008 not in(04,05,08,09,12,16,62,88)
    and V001009 not in(04,05,08,09,12,16,62,88)
    and V001012 ne 1
    then VCF0105a = 9 ;
* Race not missing or Hispanic but separate Hispanic status items missing data ;
if V001006a not in(40,00,98,99) and V001006b ne 40
    and V001006c ne 40
    and V001008 in(00,98,99)
    and V001012 in (0,8,9)
    then VCF0105a = 9 ;

go to SKIP ;
end ;
* 2002 NOTE: For the release of the race summary variable for the 2002 study, ;
* (V023150), texts of 'Other-Specify' answers have been reviewed and, if found ;
* to constitute a group mention corresponding to codes 01,02,03,04, or 05 as described ;
* above, have been 'upcoded' appropriately. For example, an "African-American" ;
* response for Other-specify belongs to category 'Black' rather than to category 'Other'. ;
* All cases except two in category 'Other' (77) represent 3 or more of major 'race' ;
* groups, while codes 12-45 are used for various combinations of 2 major group mentions. ;
* In 2002, the main ethnicity-nationality question is not asked, but for ;
* 2002 panel cases who were previously interviewed in 2000 (V021001=1) ;
* data from the 2000 variables for this question (V001008, V001009) are used ;
* in VCF0105a. ;
* The question asking whether R is of Hispanic descent is not administered to panel ;
* respondents in 2002 but is administered to nonpanel respondents who do not ;
* give a Hispanic mention for the 'race' question. For panel cases, however, the 2000 ;
* study data for the question have been included in V023151. ;
if year = 2002 then do ;
    if V023150 = 01
        then VCF0105a = 2 ;
    if V023150 = 02
        then VCF0105a = 3 ;
    if V023150 = 03
        then VCF0105a = 4 ;
    if V023150 in(04,45,14,24,34)
        then VCF0105a = 5 ;
    if V023150 = 05
        then VCF0105a = 1 ;
    if V023150 in(12,13,15,23,25,35)
        then VCF0105a = 6 ;
    if V023150 = 77
        then VCF0105a = 6 ;
    if V021001=1 then do ;
        if yr02_V001008 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
        if yr02_V001009 in(04,05,08,09,12,16,62,88) then VCF0105a = 5 ;
    end ;
if V023151 = 1 then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V021001=1 and V023150 in(88,89,99) and yr02_V001008 not in(04,05,08,09,12,16,62,88)
    and yr02_V001009 not in(04,05,08,09,12,16,62,88)
    and V023151 ne 1
    then VCF0105a = 9 ;
if V021001 ne 1 and V023150 in(88,89,99) and V023151 ne 1
    then VCF0105a = 9 ;
* Race not missing or Hispanic but separate Hispanic status items missing data ;
if V021001=1 and V023150 not in(04,45,14,24,34,88,89,99)
    and yr02_V001008 in(00,98,99)
    and V023151 in(0,8,9)
    then VCF0105a = 9 ;
if V021001 ne 1 and V023150 not in(04,45,14,24,34,88,89,99)
    and V023151 in(0,8,9)
    then VCF0105a = 9 ;

go to SKIP ;
end ;
* 2004 NOTE: In 2004, questions are the same as in 2000 (allowing 3 mentions for ;

```

MemoAppendix_race-ethrcode.txt

* ethnicity-nationality, however). The ethnicity master code is revised to a set of ;
 * 3-digit codes which closely parallel the set of 2-digit codes used in previous ;
 * years. Ethnicity-nationality mentions coded to Hispanic in VCF0105a include: ;
 * 40 Mexican, 50 Central American, 80 Cuban, 90 Dominican Republic, ;
 * 120 Puerto Rican, 160 South American, 620 Spanish, 863 Chicano or Mexican- ;
 * American.
 * For the release of the race summary variable for the 2004 study (V043299), texts of ;
 * 'Other-Specify' answers have been reviewed and, if found to constitute a group mention ;
 * corresponding to codes 10,20,30,40, or 50 as described above, have been 'upcoded' ;
 * appropriately. For example, an "African-American" response for Other-specify belongs ;
 * to category 'Black' rather than to category 'Other'. ;
 * No more than 2 mentions were given by any respondents for 'race' in 2004, and codes ;
 * 12-15 23-25,34,35,45 in V043299 are used for various combinations of 2 major group mentions. ;
 if year = 2004 then do ;
 if V043299 = 10 then VCF0105a = 2 ;
 if V043299 = 20 then VCF0105a = 3 ;
 if V043299 = 30 then VCF0105a = 4 ;
 if V043299 in(40,14,24,34,45) then VCF0105a = 5 ;
 if V043299 = 50 then VCF0105a = 1 ;
 if V043299 in(12,13,15,23,25,35,70) then VCF0105a = 6 ;
 if V043301a in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
 if V043301b in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
 if V043301c in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
 if V043305 = 1 then VCF0105a = 5 ;
 * Race missing and no Hispanic indication ;
 if V043299 in(88,89) and V043301a not in(40,50,80,90,120,160,620,863)
 and V043301b not in(40,50,80,90,120,160,620,863)
 and V043301c not in(40,50,80,90,120,160,620,863)
 and V043305 ne 1 then VCF0105a = 9 ;
 * Race not missing or Hispanic but separate Hispanic status items missing data ;
 if V043299 not in(04,14,24,34,45,88,89) and V043301a in(888,889)
 and V043305 in(0,8,9) then VCF0105a = 9 ;
 go to SKIP ;
 end ;
 * 2008 NOTE: In 2008, questions are the same as in 2000 (allowing 4 mentions for ;
 * ethnicity-nationality, however). The ethnicity master code used in 2004 is used ;
 * again. ;
 * In 2008, the question asking whether R is of Hispanic descent is asked of Rs who ;
 * did not give an Hispanic mention for either 'race' or ethnicity-nationality. ;
 * For the release variables for race mentions in 2008, texts of 'Other-Specify' ;
 * answers have been reviewed and, if found to constitute a group mention ;
 * corresponding to codes 10,20,30,40, or 50 as described above, have been re-assigned ;
 * to a special code rather than remaining as Other-specify. For example, an ;
 * "African-American" response for Other-specify is re-assigned to special category ;
 * 'Other- recoded to black' (81). ;
 * A maximum of 5 mentions are present for 'race' in 2008 (a 'race' summary variable ;
 * has not been created). ;
 if year = 2008 then do ;
 * only 1 answer ;
 if V083251b = -1 then do ;
 if V083251a in(10,81) then VCF0105a = 2 ;
 if V083251a in(20,82) then VCF0105a = 3 ;
 if V083251a in(30,83) then VCF0105a = 4 ;
 if V083251a in(40,84) then VCF0105a = 5 ;

MemoAppendix_race-ethrcode.txt

```

if V083251a in(50,85) then VCF0105a = 1 ;
if V083251a = 90 then VCF0105a = 6 ;
end ;
* more than 1 answer ;
if V083251b ne -1 then do ;
if V083251a in(40,84) or V083251b in(40,84) or V083251c in(40,84)
or V083251d in(40,84) or V083251e in(40,84) then VCF0105a = 5 ;
else do ;
black = 0 ; asian = 0 ; nativeam = 0 ; white = 0 ;
if V083251a in(10,81) or V083251b in(10,81) or V083251c in(10,81)
or V083251d in(10,81) or V083251e in(10,81)
then black = 1 ;
if V083251a in(20,82) or V083251b in(20,82) or V083251c in(20,82)
or V083251d in(20,82) or V083251e in(20,82)
then asian = 1 ;
if V083251a in(30,83) or V083251b in(30,83) or V083251d in(30,83)
or V083251d in(30,83) or V083251e in(30,83)
then nativeam = 1 ;
if V083251a in(50,85) or V083251b in(50,85) or V083251e in(50,85)
or V083251d in(50,85) or V083251e in(50,85)
then white = 1 ;
numeth = black + asian + nativeam + white ;
if numeth gt 1 then VCF0105a = 6 ;
if numeth eq 1 and black = 1 then VCF0105a = 2 ;
if numeth eq 1 and asian = 1 then VCF0105a = 3 ;
if numeth eq 1 and nativeam = 1 then VCF0105a = 4 ;
if numeth eq 1 and white = 1 then VCF0105a = 1 ;
if numeth eq 0 and (V083251a = 90 or V083251b = 90
or V083251c = 90 or V083251d = 90
or V083251e = 90) then VCF0105a = 6 ;
end ;
end ;
if V083252a in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
if V083252b in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
if V083252c in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
if V083252d in(40,50,80,90,120,160,620,863) then VCF0105a = 5 ;
if V083256 = 1 then VCF0105a = 5 ;
* Race missing and no Hispanic indication ;
if V083251a in(-8,-9) and V083252a not in(40,50,80,90,120,160,620,863)
and V083252b not in(40,50,80,90,120,160,620,863)
and V083252c not in(40,50,80,90,120,160,620,863)
and V083252d not in(40,50,80,90,120,160,620,863)
and V083256 ne 1 then VCF0105a = 9 ;
* Race not missing or Hispanic but separate Hispanic status items missing data ;
if V083251a not in(40,84,-8,-9) and V083251b not in(40,84)
and V083251c not in(40,84)
and V083251d not in(40,84)
and V083251e not in(40,84)
and V083252a in(-8,-9,-4)
and V083256 in (-8,-9,-99) then VCF0105a = 9 ;
go to SKIP ;
end ;
* ;
* ----- ;

```

MemoAppendix_race-ethrcode.txt

- * In 2012, race-ethnicity questions follow a new sequence. ;
- * Respondents are first asked 'Are you Spanish, Hispanic or Latino?' (Yes/No).
- * The race question is asked next and has a new format ('I am going to read you ;
- * a list of five race categories. Please choose one or more races that you consider ;
- * yourself to be') which provides categories from the Census Bureau ;
- * White (01), Black or African-American (02), American Indian or Alaska Native ;
- * (03), Asian (04), Native Hawaiian or other Pacific Islander (05).
- * The 1988-2008 question asking whether R is of Hispanic descent is not present. ;
- * Although the open-end ethnicity-nationality question asked since 1972 is ;
- * administered again in 2012, it is no longer used for VCF0105a. ;
- * Cases missing in the CDF summary VCF0105a are those having: ;
- * 1) missing data for race together with Hispanic status not indicated (the ;
- * question asking if R is Spanish, Hispanic or Latino is missing or No), or ;
- * 2) race is not nonresponse but there is missing data for the question asking ;
- * if R is Spanish, Hispanic or Latino. ;
- * 2012 NOTE: ;
- * For the 2012 study, races other than black and white were originally restricted ;
- * in early releases, and the 2012 race-and-ethnicity summary variable (dem_raceeth_x) ;
- * was created with categories only for 'Hispanic', 'White non-Hispanic', ;
- * 'Black non-Hispanic', and 'Other non_Hispanic.' ;
- * However, due to requests from the user community, codes providing greater specificity ;
- * for individual race mentions have since been added to the study release. ;
- * Unfortunately, the race-and-ethnicity summary variable itself has not yet been revised ;
- * accordingly in order to represent the fuller disclosure in the individual mentions ;
- * of race. ;
- * For the addition of 2012 data to VCF0105a, the 2012 summary variable for race ;
- * must first be updated, and the same update will be applied in the next release of ;
- * the 2012 study. ;
- * New codes for 2012 study summary var dem_raceeth_x are the following: ;
- * 1 White, non-Hispanic ;
- * 2 Black, non-Hispanic ;
- * 3 Asian, native Hawaiian or other Pacific Islander, non-Hispanic ;
- * 4 Native American or Alaska Native, non-Hispanic ;
- * 5 Hispanic ;
- * 6 Other, including multiple races ;
- * In addition, codes for variable dem_racecps_othrace are changed to the following in order ;
- * to upcode Other-specify texts according to the greater disclosure: ;
- * 00 Not selected ;
- * 01 Selected, not blank and not codeable to one category among: White, Black, Asian/native Hawaiian/Pacific Islander, ;
- * Native American/Alaska native, Hispanic ;
- * 10 Other-specify text codeable to White ;
- * 20 Other-specify text codeable to Black ;
- * 30 Other-specify text codeable to Asian ;
- * 35 Other-specify text codeable to Pacific Islander ;
- * 40 Other-specify text codeable to Native American or Alaska Native ;
- * 50 Other-specify text codeable to Hispanic ;
- * 90 WEB ONLY: Selected by R, text is blank ;
- * It should be noted that the choice 'Other' was not read to FTF Rs but was a listed response option for Web respondents. ;
- * Upcoding of Other-specify must occur first. ;
- if year = 2012 then do ;
- if caseid
- in(727,755,913,1051,1298,1748,3052,3203,3215,3358,4428,4434,4458,4589,4686,4737,4904,5029,5240,5479,5806,583

MemoAppendix_race-ethrcode.txt

```

7,5842,6328,6556,6833)
  then dem_racecps_othrace = 10 ;
  if caseid in(955,1309,1694,3140,3330,4096,6302)
    then dem_racecps_othrace = 20 ;
  if caseid in(535,809,1877,5224)
    then dem_racecps_othrace = 40 ;
  if caseid
in(129,310,318,472,492,493,504,519,524,525,536,623,630,632,633,670,672,674,675,677,685,690,693,694,695,698,708,
729,730,771,798,834,857,866,867,1126,1148,1300,1393,1409,1410,1411,1412,1433,1458,1472,1505,1506,1508,1594)
  then dem_racecps_othrace = 50 ;
  if caseid
in(1658,1771,1814,1816,1824,1832,1833,1863,1880,1893,1894,1924,1945,1953,1992,2035,3093,3247,3309,3412,3440,
3795,3902,3966,4167,4284,4448,4467,4479,4490,4515,4564,4611,4621,4627,4833,4834,4869,4887,4919,4924,4947)
  then dem_racecps_othrace = 50 ;
  if caseid
in(5038,5055,5077,5120,5132,5179,5180,5234,5259,5290,5316,5416,5433,5509,5519,5604,5630,5690,5904,6063,6234,
6246,6252,6263,6272,6337,6445,6512,6578,6586,6644,6650,6666,6719,6785,6788,6817,6834,6838,6841,6863,5710)
  then dem_racecps_othrace = 50 ;
  if caseid
in(3078,3251,3353,3529,3841,3943,4271,4561,4847,4879,4922,4923,4962,5237,5436,5704,5727,5755,5757,5822,5845,
5849,5941,6002,6035,6050,6170,6178,6179,6185,6194,6253,6310,6321,6330,6342,6388,6406,6617,6645,6658,6684,671
7,6736,6764,6776,6800,6847)
  then dem_racecps_othrace = 90 ;
  numrace = 0 ;
  black = 0 ; asian = 0 ; nativeam = 0 ; white = 0 ;
  if dem_racecps_white = 1 or dem_racecps_othrace = 10 then white = 1 ;
  if dem_racecps_black = 1 or dem_racecps_othrace = 20 then black = 1 ;
  if dem_racecps_asian = 1 or dem_racecps_pacif = 1 or dem_racecps_othrace in(30,35) then asian = 1 ;
  if dem_racecps_native = 1 or dem_racecps_othrace = 40 then nativeam = 1 ;
  numrace = black + asian + nativeam + white ;
  if numrace gt 1 then VCF0105a = 6 ;
  if numrace eq 1 and black = 1 then VCF0105a = 2 ;
  if numrace eq 1 and asian = 1 then VCF0105a = 3 ;
  if numrace eq 1 and nativeam = 1 then VCF0105a = 4 ;
  if numrace eq 1 and white = 1 then VCF0105a = 1 ;
  if numrace eq 0 and dem_racecps_othrace in(1,90) then VCF0105a = 6 ;
  if dem_racecps_racedkrf in(1,2) then VCF0105a = 9 ;
  if dem_racecps_othrace eq 50 then VCF0105a = 5 ;
  if dem_hisp eq 1 then VCF0105a = 5 ;
  go to SKIP ;
end ;
MD:
  if VCF0105a = . then VCF0105a = 9 ;
SKIP:
run;

proc formats library=library ;
VALUE VCF0105a
  1 = "1. White non-Hispanic (1948-2012)"
  2 = "2. Black non-Hispanic (1948-2012)"
  3 = "3. Asian or Pacific Islander, non-Hispanic (1966-2012)"
  4 = "4. American Indian or Alaska Native non-Hispanic (1966-2012)"
  5 = "5. Hispanic (1966-2012)"
  6 = "6. Other or multiple races, non-Hispanic (1968-2012)"

```

```

7 = "7. Non-white and non-black (1948-1964)"
9 = "9. Missing" ;
VALUE VCF0106_
1 = "1. White non-Hispanic"
2 = "2. Black non-Hispanic"
3 = "3. Other"
0 = "0. Missing, pre-1966 data"
9 = "9. Missing, DK/REF/NA";
VALUE VCF0106a
1 = "1. White non-Hispanic"
2 = "2. Black non-Hispanic"
3 = "3. Hispanic"
4 = "4. Other or multiple races, non-Hispanic"
0 = "0. Missing, pre-1966 data"
9 = "9. Missing, DK/REF/NA";
run;

data library.cdf_newvars ;
  set library.cdf_raceeth (keep = year caseid VCF0105a);
* added summary VCF0106 ;
if VCF0105a = 1 then VCF0106 = 1 ;
if VCF0105a = 2 then VCF0106 = 2 ;
if VCF0105a = 3 then VCF0106 = 3 ;
if VCF0105a = 4 then VCF0106 = 3 ;
if VCF0105a = 5 then VCF0106 = 3 ;
if VCF0105a = 6 then VCF0106 = 3 ;
if VCF0105a = 7 then VCF0106 = 0 ;
if VCF0105a = 9 then VCF0106 = 9 ;
* new summary VCF0106a ;
if VCF0105a = 1 then VCF0106a = 1 ;
if VCF0105a = 2 then VCF0106a = 2 ;
if VCF0105a = 3 then VCF0106a = 4 ;
if VCF0105a = 4 then VCF0106a = 4 ;
if VCF0105a = 5 then VCF0106a = 3 ;
if VCF0105a = 6 then VCF0106a = 4 ;
if VCF0105a = 7 then VCF0106a = 0 ;
if VCF0105a = 9 then VCF0106a = 9 ;
* ;
FORMAT VCF0105a VCF0105a. VCF0106 VCF0106_. VCF0106a VCF0106a. ;
run ;

```