

**Fall Enrollment Data File, 1997  
Survey Methodology and  
User Guidelines**

## Table of Contents

<b>Survey Methodology .....</b>	<b>2</b>
Overview .....	2
Universe, Institutions Surveyed and Response Rates .....	4
Survey Conduct and Editing .....	7
Data Management and Imputation .....	8
Raking of Racial-ethnic Data .....	8
<b>User Guidelines for Processing the 1997 IPEDS Fall Enrollment Data File.....</b>	<b>10</b>
Overview .....	10
General Information .....	10
File Structure .....	12
Data Analysis: Subsetting the Records .....	14
Cautionary Notes .....	15

### List of Tables

Table A.	Number of postsecondary institutions in the 1997 Fall Enrollment survey and survey response rates, by survey form and sector of institution: 50 states and the District of Columbia and the outlying areas .....	5
Table B.	Number and survey response rates of postsecondary institutions, by Title IV eligibility, degree-granting status, and sector of institution: 50 states and the District of Columbia and the outlying areas .....	6
Table C.	Total enrollment by race-ethnicity unknown and balance, by Title IV eligibility, degree-granting status, and sector of institution: 50 states and the District of Columbia and the outlying areas .....	9

# Survey Methodology

## Overview

The Fall Enrollment (EF) survey is conducted annually as part of the National Center for Education Statistics' Integrated Postsecondary Education Data System (IPEDS).<sup>1</sup> The EF survey collects data on the number of students enrolled in postsecondary institutions<sup>2</sup> that were eligible to participate in Title IV federal financial aid programs in the United States and its outlying areas<sup>3</sup>. Eligibility status of institutions was obtained from the Office of Postsecondary Education's March 1997 Postsecondary Education Participants System (PEPS) file.

Students **included** in the Fall Enrollment survey:

- students enrolled in courses creditable toward a degree or other formal award;
- students enrolled in courses that are part of a vocational or occupational program, including those enrolled in off-campus centers; and
- high school students taking regular college courses for credit.

Students **excluded** from the Fall Enrollment survey:

- students enrolled exclusively in courses not creditable toward a formal award and who are not in a postsecondary vocational program;
- students enrolled exclusively in remedial courses;
- students exclusively auditing classes;
- students studying abroad (e.g., at a foreign university) if their enrollment at the institution is only an administrative record and the fee is only nominal;
- students in any branch campus located in a foreign country; and
- students earning continuing education units (CEU's) only.

---

<sup>1</sup> Other IPEDS surveys include Institutional Characteristics, Completions, Finance, Faculty Salaries, Graduation Rates, Fall Staff, Academic Libraries, and the Consolidated survey. NCEs conducts IPEDS surveys in order to fulfill its legislative mission "to collect, and analyze, and disseminate statistics and other data related to education in the United States" (P.L. 103-382, National Education Statistics Act of 1994, Sec. 404(a)).

<sup>2</sup> Postsecondary education is defined as the provision of a formal instructional program whose curriculum is designed primarily for students who are beyond the compulsory age for high school. This includes programs whose purpose is academic, vocational, and continuing professional education, and excludes avocational (leisure) and adult basic education programs.

<sup>3</sup> The outlying areas include the Federated States of Micronesia, Guam, the Marshall Islands, Northern Marianas, Palau, Puerto Rico, and the Virgin Islands.

There are two versions of the Fall Enrollment survey. The most extensive form, "EF1", was sent to all 4-year institutions. The "EF2" form is less detailed and was sent to 2-year postsecondary institutions that grant an associate's degree (degree-granting). Additional enrollment data was collected on the "Consolidated" survey (CN), which was sent to 2-year institutions that grant awards or certificates of at least 2 but less than 4 academic years (non-degree-granting) and less-than-2-year institutions that offer awards or certificates of less than 2 years duration<sup>4</sup>. The enrollment data collected via the Consolidated survey are integrated into the Fall Enrollment database. The data requested on each form or survey varies as follows:

Form Type	Level of Institution	Data Requested:
EF1	4-year institutions	Attendance Status (full- or part-time) Student Level (undergraduate, first-professional, graduate) Race-ethnicity and sex of student Degree-seeking status Major field of study (even numbered years only)  Year of study Age of student (odd numbered years only) Residence of first-time students (even numbered years only)
EF2	2-year degree-granting	Attendance Status (full- or part-time) Race-ethnicity and sex of student Degree-seeking status Year of study Age of student (odd numbered years only) Residence of first-time students (even numbered years only)
CN	All other institutions	Attendance Status (full- or part-time) Race-ethnicity and sex of student Year of study

---

<sup>4</sup> The Consolidated survey is designed to collect information from Title IV eligible, non-degree-granting institutions on fall enrollment, finance, fall staff, and completions.

## **Universe, Institutions Surveyed and Response Rates**

A major subset of all postsecondary institutions are those institutions that are eligible to participate in Title IV federal financial aid programs. It is this group of institutions that annually receive a fall enrollment survey. Beginning with the 1996 IPEDS, the subset of eligible institutions was validated by matching the IPEDS universe with the PEPS file, which is maintained by the Office of Postsecondary Education (OPE). OPE grants eligibility to institutions to participate in Title IV federal financial programs. This validation process was repeated for the 1997 IPEDS for determining eligibility, and 6,810 eligible institutions were mailed Fall Enrollment or Consolidated Survey forms.

Survey forms were mailed in July 1997: (1) EF1 forms were mailed to 2,365 institutions; (2) EF2 forms to 1,772 institutions; and (3) CN surveys to 2,673 institutions. The survey results were collected from November 1997 through June 1998. During this time period a total of 175 institutions were deleted from the universe. These deletions resulted from formal notification by IPEDS state coordinators, the Department of Education eligibility notices, and from follow-up telephone calls. Included in the deletions were: (1) duplicates of other institutions on the file; (2) institutions that closed or had merged with another institution, and thus were no longer a legitimate institution or branch; (3) institutions that no longer offered postsecondary programs; or (4) schools that did not conform to the IPEDS definition of an institution or branch. At the conclusion of this process, 6,635 institutions represented the final 1997 fall enrollment universe. The final universe was also adjusted to reflect institutions that changed from one sector to another subsequent to survey mailout.

Table A shows the number of institutions that were mailed the 1997 Fall Enrollment survey (or the Consolidated form) by level and control of institution; the final in-scope universe of institutions receiving either survey form; the number of in-scope institutions that responded; and the response rates.

**Table A. –Number of postsecondary institutions in the 1997 Fall Enrollment survey and survey response rates, by survey form and sector of institution: 50 states, District of Columbia, and the outlying areas**

Survey form and sector of institution	Mailed	Final universe	Number responded	Response rate
All postsecondary	6,810	6,635	6,285	94.7
EF1 form	2,365	2,362	2,270	96.1
EF2 form	1,772	1,743	1,671	95.9
Consolidated survey	2,673	2,530	2,344	92.6
4-year	2,365	2,382	2,289	96.1
Public	631	634	626	98.7
Private non-profit	1,581	1,583	1,503	94.9
Private for-profit	153	165	160	97.0
2-year	2,494	2,414	2,311	95.7
Public	1,248	1,220	1,195	98.0
Private non-profit	390	362	335	92.5
Private for-profit	856	832	781	93.9
Less-than-2-year	1,951	1,839	1,685	91.6
Public	215	224	214	95.5
Private non-profit	100	93	88	94.6
Private for-profit	1,636	1,522	1,383	90.9

NOTE: By definition, 4-year institutions offer a bachelor's degree or postbaccalaureate award of some kind; 2-year institutions offer at least one program of at least 2 years in duration; and less-than-2-year schools offer only programs of less-than-2-years in duration. The sector of an institution at the time of mailout may differ from its sector in the final universe because a school's sector may change due to a shift in the highest level of offering or due to improper classification by sector the prior year. Because of this sector change some 4-year institutions may return an EF2 form and some 2-year institutions may return an EF1 form. Response rates were calculated as the ratio of the number of completed survey forms divided by the number of schools in the final universe.

SOURCE: U.S Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System, "Fall Enrollment" survey 1997.

On the 1997 IPEDS files, institutions are identified as (1) degree-granting - those that offer either an associate's, bachelor's, master's, doctor's, or first-professional degree; and (2) non-degree-granting (certificate only) - those institutions that offer only certificates at any level. Institutions offering only postbaccalaureate, post-master's, or first-professional certificates and institutions offering certificates of less-than-4-years are included in this group.

Table B provides the number of institutions and survey response rates by eligibility status and degree-granting status by control and level of institution. The institution's eligibility status was updated in March of 1998 and 246 institutions that were sent a fall enrollment or consolidated survey lost eligibility status. There are 6,389 eligible institutions on the file; of these, 4,074 are degree-granting.

**Table B. –Number and survey response rates of postsecondary institutions, by Title IV eligibility, degree-granting status, and sector of institution: 50 states, District of Columbia, and the outlying areas, Fall 1997**

Degree-granting status and sector of institution	All institutions			Eligible			Not eligible		
	Final universe	Number responded	Response rate	Final universe	Number responded	Response rate	Final universe	Number responded	Response rate
All institutions	6,635	6,285	94.7	6,389	6,122	95.8	246	163	66.3
4-year									
Public	634	626	98.7	633	625	98.7	1	1	100.0
Private non-profit	1,583	1,503	94.9	1,563	1,491	95.4	20	12	60.0
Private for-profit	165	160	97.0	162	157	96.9	3	3	100.0
2-year									
Public	1,220	1,195	98.0	1,218	1,193	97.9	2	2	100.0
Private non-profit	362	335	92.5	333	317	95.2	29	18	62.1
Private for-profit	832	781	93.9	799	757	94.7	33	24	72.7
Less-than-2-year									
Public	24	214	95.5	217	209	96.3	7	5	71.4
Private non-profit	93	88	94.6	88	85	96.6	5	3	60.0
Private for-profit	1,522	1,383	90.9	1,376	1,288	93.6	146	95	65.1
Degree granting	4,109	3,947	96.1	4,074	3,925	96.3	35	22	62.9
4-year									
Public	629	621	98.7	629	621	98.7	0	0	—
Private non-profit	1,549	1,473	95.1	1,538	1,468	95.4	11	5	45.5
Private for-profit	160	156	97.5	157	153	97.5	3	3	100.0
2-year									
Public	1,099	1,076	97.9	1,099	1,076	97.9	0	0	—
Private non-profit	179	167	93.3	173	165	95.4	6	2	33.3
Private for-profit	493	454	92.1	478	442	92.5	15	12	80.0
Non-Degree granting	2,526	2,338	92.6	2,315	2,197	94.9	211	141	66.8
4-year									
Public	5	5	100.0	4	4	100.0	1	1	100.0
Private non-profit	34	30	88.2	25	23	92.0	9	7	77.8
Private for-profit	5	4	80.0	5	4	80.0	0	0	—
2-year									
Public	121	119	98.3	119	117	98.3	2	2	100.0
Private non-profit	183	168	91.8	160	152	95.0	23	16	69.6
Private for-profit	339	327	96.5	321	315	98.1	18	12	66.7
Less-than-2-year									
Public	224	214	95.5	217	209	96.3	7	5	71.4
Private non-profit	93	88	94.6	88	85	96.6	5	3	60.0
Private for-profit	1,522	1,383	90.9	1,376	1,288	93.6	146	95	65.1
50 states and D.C.	6,492	6,169	95.0	6,252	6,010	96.1	240	159	66.3
Degree granting	4,032	3,882	96.3	3,998	3,861	96.6	34	21	61.8
Non-degree granting	2,460	2,287	93.0	2,254	2,149	95.3	206	138	67.0
Outlying areas	143	116	81.1	137	112	81.8	6	4	66.7
Degree granting	77	65	84.4	76	64	84.2	1	1	100.0
Non-degree granting	66	51	77.3	61	48	78.7	5	3	60.0

— not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System, "Fall Enrollment" survey 1997.

## Survey Conduct and Editing

The 1997 Fall Enrollment survey was due November 15. Survey data were collected via paper or transmitted to NCES via the Internet. Some institutions did submit data on diskette. Data for nonresponding CN institutions were collected through the Postsecondary Education Telephone System (PETS). All data, whether received on paper forms, diskettes, electronically via the Internet, or through the PETS system, went through the same editing process. Extensive follow-up for survey nonresponse was conducted from November 1997 through April 1998. Initially, reminder letters were mailed, encouraging nonresponding institutions to complete and return their forms; and subsequently, the PETS was used to collect critical data by telephone from an institutional representative.

Survey responses were edited for internal and inter-year consistency. The following editing procedures were used:

### Part A: Enrollment by sex and race-ethnicity

- (1) Addition checks were performed by adding down the columns and comparing generated totals with reported totals. If the reported total differed from the generated total but was within a designated range, the reported total was replaced by the generated total and the cell was flagged with the proper impute code. If the difference exceeded the designated range, institutions were contacted for verification/correction. Addition checks were also performed by generating totals for men and women by adding across columns (racial-ethnic group) in each line. Generated totals by sex were compared to the reported totals. If they differed but were within a designated range a balance field was created containing the difference by sex and placed on the data file. Reported detail and totals were not altered. If the difference exceeded the designated range, institutions were contacted for correction. Editing of the racial-ethnic data is explained in detail in the section on raking of racial-ethnic data.
- (2) Comparisons were also made with the number of students reported for selected items for the prior year. If the differences were sufficiently large to trigger an edit flag, institutions were contacted for verification.

### Part B: Enrollment by Age

- (1) Part B data were edited in a similar manner. Addition checks were performed by comparing reported totals in Part B to corresponding totals in Part A. Differences between the totals were calculated and balances containing the differences were generated. If the balance was outside a certain range the institution was contacted for verification/correction. If the balance was within a certain range, the Part B total was replaced by the corresponding total from Part A. The balance was then allocated among the age categories using the institution's reported age distribution.

## **Data Management and Imputation**

Data for Part A, enrollment by race-ethnicity, were imputed for nonresponding institutions. The following imputation procedures were used:

For institutions with reported data on the prior year enrollment file: all institutions that reported data in both 1996 and 1997 were grouped by level of offering and control of institution. Enrollment data were then aggregated for each group and percent changes were calculated. For each nonresponding institution in 1997, their 1996 enrollment data were adjusted by the percent change of the group to which that institution belonged.

For institutions with no prior year report, imputations were based on current year group means or ratios derived for similar institutions with reported data in 1997. Groups were formed by combining institutions by level of offering and control of institution.

## **Raking of Racial-Ethnic Data**

When data were reported by racial-ethnic categories on a detail line and the generated sum of these enrollments did not equal the reported total enrollment, the difference between the generated total and the reported total was calculated. If the difference exceeded a certain designated range the institution was contacted for correction. If the difference was within the designated range, a "balance column" was created. The balance was then distributed in the same proportions as the reported racial-ethnic data for that detail line. If there were no racial-ethnic data reported for any detail lines, but racial-ethnic data were reported for the subtotal lines (8=undergraduate full-time, 14=postbaccalaureate full-time, 22=undergraduate part-time and 28=postbaccalaureate part-time), the total from columns 15 and 16 (male, female ) of each detail line were distributed in the same proportions as the racial-ethnic enrollments reported on the corresponding subtotal line. If the institution did not report racial-ethnic data on the subtotal lines, but reported racial-ethnic data on the grand total line (29), then line 29 was used to distribute the data on the detail lines. If the institution only reported data in columns 15 and 16 on all lines, data were distributed among racial-ethnic categories based on the distribution of totals of the imputation group in which the institution was a member. Groups were formed by aggregating lines 8, 14, 22, and 28 by survey form, control, and state by sex.

Beginning in 1990, a "race-ethnicity unknown category" was incorporated into the Fall Enrollment survey. To allow for analysis of enrollment trends by race-ethnicity, enrollment reported in the race-ethnicity unknown category was raked across those columns that represent U.S. citizens, that is, columns 3 through 12; race unknown was not distributed to nonresident aliens. The procedure used to distribute the race unknown category among the race-ethnic categories was the same procedure used for distributing the balance columns. Note that raking may be done on data that are imputed as well as the reported data. Table C contains the enrollment numbers by "race-ethnicity unknown" and balance categories.

**Table C. –Total enrollment by race-ethnicity unknown and balance, by Title IV eligibility, degree-granting status, and sector of institution: 50 states, District of Columbia, and the outlying areas, Fall 1997**

Degree-granting Status and sector of institution	All institutions			Eligible			Not eligible		
	Total enrollment	Percent Race-ethnicity Unknown	Percent Balance	Total enrollment	Percent race-ethnicity unknown	Percent Balance	Total enrollment	Percent race-ethnicity unknown	Percent balance
All institutions	15,133,663	4.0	—	15,098,291	4.0	—	35,372	17.2	1.2
4-year									
Public	5,908,380	3.0	—	5,908,377	3.0	—	3	0.0	0.0
Private non-profit	2,995,436	5.7	—	2,991,216	5.7	—	4,220	28.8	0.0
Private for-profit	161,062	5.0	—	159,044	4.9	—	2,018	11.3	0.0
2-year									
Public	5,465,244	3.7	—	5,465,153	3.7	—	91	0.0	0.0
Private non-profit	87,468	3.8	—	85,921	3.8	—	1,547	2.8	0.0
Private for-profit	232,760	7.7	—	221,197	6.2	—	11,563	35.8	0.0
Less-than-2-year									
Public	80,506	9.6	—	77,940	9.9	—	2,566	0.7	0.0
Private non-profit	23,077	48.5	0.0	22,681	49.4	0.0	396	0.0	0.0
Private for-profit	179,730	3.6	0.2	166,762	3.6	0.0	12,968	3.5	3.3
Degree granting	14,697,956	3.9	—	14,680,488	3.9	—	17,468	31.9	0.0
4-year									
Public	5,908,293	3.0	—	5,908,293	3.0	—	—	—	—
Private non-profit	2,993,859	5.7	—	2,989,830	5.7	—	4,029	30.1	0.0
Private for-profit	160,644	5.0	—	158,626	4.9	—	2,018	11.3	0.0
2-year									
Public	5,368,541	3.7	—	5,368,541	3.7	—	—	—	—
Private non-profit	73,957	4.0	0.0	72,956	4.0	0.0	1,001	1.4	0.0
Private for-profit	192,662	6.0	—	182,242	4.1	—	10,420	39.5	0.0
Non-degree granting	435,707	7.7	0.1	417,803	7.9	—	17,904	2.9	2.4
4-year									
Public	87	0.0	0.0	84	0.0	0.0	3	0.0	0.0
Private non-profit	1,577	0.5	0.0	1,386	0.5	0.0	191	0.5	0.0
Private for-profit	418	4.8	0.0	418	4.8	0.0	—	—	—
2-year									
Public	96,703	1.5	—	96,612	1.5	—	91	0.0	0.0
Private non-profit	13,511	2.6	0.1	12,965	2.5	0.1	546	5.3	0.0
Private for-profit	40,098	15.8	0.0	38,955	16.2	0.0	1,143	1.3	0.0
Less-than-2-year									
Public	80,506	9.6	—	77,940	9.9	—	2,566	0.7	0.0
Private non-profit	23,077	48.5	0.0	22,681	49.4	0.0	396	0.0	0.0
Private for-profit	179,730	3.6	0.2	166,762	3.6	0.0	12,968	3.5	3.3

— less than 0.05 percent.

NOTE: Data in table J were compiled using only line 29 (total enrollment). Balance columns were generated when the calculated total enrollment by race-ethnicity did not match the reported total. The balance is the difference between reported and calculated totals. The enrollments in the balance column arise from 128 institutions whose calculated enrollments by race-ethnicity were not equal to the reported totals. Two of these institutions reported no race-ethnicity data and did not report data in the race-ethnicity unknown category. Enrollments imputed from allocating the balance columns is below 0.05 percent. Enrollment in the "race-ethnicity unknown" category was reported by 2,673 institutions. Sixty-nine of these institutions reported data only in the race-ethnicity unknown category.

SOURCE: U.S Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System, "Fall Enrollment" survey 1997.

## User Guidelines for Processing the 1997 IPEDS Enrollment Data File

### Overview

The 1997 Fall Enrollment and Consolidated surveys, part of the Integrated Postsecondary Education Data System (IPEDS), provide a vehicle for estimating enrollments by race/ethnicity, sex, level and status for all postsecondary education providers in the IPEDS universe. The file contains fall enrollment data for 4-year, 2-year, less-than-2-year institutions. This information can be used to answer such questions as:

- a. What is the total number of full-time and part-time students attending the nation's postsecondary institutions?
- b. What proportion of degree seeking students are in graduate programs?
- c. What proportion of degree seeking students attend school on a part-time basis? Does this vary by the level of institution?

### General Information

The following general guidelines will be useful in working with the Fall Enrollment file:

1. To aid the analysis of the data file on the PC, the file has been divided into five different data file modules. These files are fixed length formatted text files.

File EF97\_IC contains the institutional characteristic information of the institution. Variables such as control, level, state fips, and abbreviation, highest level of offering and type of form the institution returned are found on this file.

File EF97\_ANR contains (Part A) summary enrollment data by reported race-ethnicity and sex of student and by level of student.

File EF97\_ARK (raked analysis Part A file) summary enrollment data by adjusted race-ethnicity and sex of student and by level of student. This file contains adjusted race-ethnicity data. The race/unknown categories and balance columns are distributed across the race-ethnicity categories. (See methodology). NCES publications are based on this file.

File EF97\_B contains (Part B) enrollment data by age categories and attendance status, and by level and sex of student.

File EF97\_D contains (Part D) clarifying questions.

2. Each institution's data consists of several records across five data file modules. Each file consists of fixed length (EF97\_IC=136 positions, EF97\_ANR=136 positions, EF97\_ARK=136 positions, EF97\_B=68 positions and EF97\_D contains 169 positions) records. The record format is the same for each institution within each file, regardless of the survey form type completed by the school. EF97\_ANR, EF97\_ARK, and EF97\_B contain

multiple records per institution and EF97\_IC and EF97\_D contain one record per institution. Storing the fall enrollment data in 5 different file modules as opposed to one big file reduces the file size about 70 percent. Recognizing that the analysis of the data in this format may be more difficult, a sample analysis using PC SAS is provided in this section.

3. Survey data items vary by form type. While some items are common to EF1, EF2, and CN, other items are unique to a specific survey form. To accurately subset the file for data analysis and tabulation, the file user should become acquainted with the data items included on each version of the survey and with the types of institutions that complete each version.
4. A UNITID number is present on every record in all data files, and is the unique identification code assigned by NCES to separately identify each postsecondary institution and administrative unit on the file. The UNITID is a constant number and will not change over time. The UNITID is used to link all enrollment data file modules together. The UNITID also links all IPEDS survey data bases for that institution (specifically, the Institutional Characteristics, Fall Enrollment, Completions, Finance, Salaries, Fall Staff, and Consolidated surveys).
5. The Office of Postsecondary Education Identification Number (OPEID) is the identification number assigned to institutions that are in the Postsecondary Education Participants System (PEPS). This system is maintained by the Office of Postsecondary Education (OPE) which tracks the eligibility status of postsecondary institutions. An institution that has an OPEID identifies it as being in both IPEDS and PEPS. The eligibility status of an institution can be found in an OPE indicator variable called (OPEIND, field position 130). All institutions that are eligible to participate in Title IV aid programs will have a value of "1" in this position. A value of "2" was given to any branch campus, or system office in IPEDS that was not identified as being eligible by PEPS but was part of a main campus that is eligible. Also, any U.S. service school not in PEPS was given a value of "2". IPEDS will consider any institution with a value of "1" or "2" as being eligible to participate in Title IV federal financial aid programs. The OPEID is an 8-digit code and is found in positions 122-129 of the data file. The first 6 digits of the OPEID identify the main campus. The last 2 digits represent branch campuses or additional locations and are coded as 01,02...99. For the main campus the last 2-digits will always be "00". For branch campuses of an eligible main campus in IPEDS that is not in PEPS, the last 2 digits are coded as A1, A2...A9.
6. As mentioned earlier IPEDS will now maintain a subset of institutions called degree-granting institutions. A degree-granting institution is any institution that offers an associate's, bachelor's, master's, doctor's, or first-professional degree. Institutions that grant only certificates or awards of any length program (less-than-2-years or 2-years or more) are categorized as non-degree-granting institutions. Codes for identifying degree-granting institutions are defined in the variable called HDEGOFFR (highest degree offered located record positions 131-132, see appendix A, list 9 for values).
7. A school's sector may change from the previous year based on its response to specific survey items. (Sector is a generated variable, and is based on an institution's control and highest level of offering.) Thus, some sector changes will lead to an apparent but correct inconsistency between form type and sector.

For example, a 4-year institution may be on the file as having completed an EF2 form, although 4-year schools are supposed to complete EF1 forms. To illustrate, let's say that at the time of mailout, "ABC College" was classified as a public 2-year school (sector = 4), and thus received and completed an EF2 form (FORMRT = 2). However, their response to the current-year Institutional Characteristics survey indicates that the institution now offers bachelor's degrees. The institution's control is still public (i.e., CONTROL = 1), but its level is now 4-year (i.e., LEVEL = 1), and its sector is now public 4-year (i.e., SECTOR = 1). Therefore, in the next survey mailout, the school will receive an EF1 form and the corresponding IPEDS forms for 4-year institutions (C1, IC1, etc.).

8. Generated total enrollment by age category aggregated from part B data will not equal generated total enrollment aggregated from Part A data, because enrollment by age were not collected on the CN form.
9. Parent/child institutions. Some institutions do not report enrollment data for individual campuses, but report data with the main campus. Institutions that report their data with another institution are known as child institutions and are identified on the file by the variable PARCHILD=2 (file position 91). Institutions that report the combined data (parent institutions) are identified by PARCHILD=1. The parent and children institutions have a data field called UNITIDX, which can be found in file positions 92-97 of the IC data module. This field contains the UNITID of the parent institution record. This data field will enable the user to link the child record with the parent record.

## **File Structure**

The file structure and survey forms of EF1, EF2, and CN should be reviewed carefully for the relative line and column positions of specific variables before any programming is started. Familiarity with the differences in EF1, EF2, and CN is especially important if a programmer's objective is to retrieve comparable data for 4-year, 2-year, and less-than-2-year institutions, or to derive national totals. A crosswalk is provided in Appendix B.

The line numbers found in record positions 8-9 correspond to the line numbers on Part A, Part B, or Part C (Note: the CN survey does not have Part B). Each line of data on the survey form corresponds to one record. In contrast, Part D of the survey is treated as one record (Note: the CN survey does not have Part D). The part code is found in record position 7. The data on each record appear in the same order as presented on the survey form. All records are of fixed length within each file module.

The variables found on a specific record and the total number of records are not necessarily comparable across EF1, EF2, and CN. For example, Part A of EF1 has 29 records of data. Each record contains data on the number of men and women, by race/ethnicity. Note there are separate records to differentiate between degree-seeking status, attendance status, and year of study. Part A of CN contains 7 records of data differentiating first-time and all other students and full-time vs. part-time. Each record contains data on the number of men and women by race/ethnicity.

Note, however, that the line numbers for EF1, EF2 and CN are consistent. For example, the grand total and subtotal line numbers for EF1, EF2, CN are the same, although 29 lines of data were collected via the EF1 form, 11 lines of data were collected via the EF2 form, and 5 lines were collected via the CN survey.

### **Structure for part A data files, total enrollment by race-ethnicity: EF97\_ANR and EF97\_ARK**

For institutions that returned an EF1 survey form, this file consists of a maximum of 29 records per institution. Each record contains enrollment data by race-ethnicity and sex of student. The record number defines the level of enrollment. Records 1 through 8 refer to full-time undergraduates by year of study. Records 9 through 14 refer to full-time post-baccalaureate students (first-professional or graduate). Records 15 through 22 refer to part-time undergraduates by year of study. Records 23 through 28 refer to part-time post-baccalaureate students (first-professional or graduate). The grand total enrollment is given on record 29, and subtotal enrollments for full-time undergraduates and postbaccalaureates and for part-time undergraduates and postbaccalaureates are given on records 8, 14, 22, and 28 respectively.

For institutions that returned an EF2 survey form, this file consists of a maximum of 11 records per institution. Each record contains enrollment data by race/ethnicity and sex of student. Records 1 through 3 refer to full-time degree seeking undergraduates by year of study. Record 7 refers to all other full-time students. Records 15 through 17 refer to part-time degree seeking undergraduates by year of study. Record 21 refers to all other part-time students. The grand total enrollment is given on record 29, full- and part-time subtotals are given on records 8 and 22. Record numbers 4 through 6, 9 through 14, 18 through 20, and 23 through 27 were not used on the EF2 form to provide consistency with similar data provided on the EF1 form.

For institutions that returned the CN survey form, this file consists of a maximum of 7 records per institution. Subtotal lines were not collected on the CN survey but were generated for the data file. (see CN/EF crosswalk in appendix B for more detail).

### **Structure of part B (total enrollment by age/categories) data file: EF97\_B**

For EF1 institutions, this file contains enrollment data by age of student. These data are collected in odd-numbered years. There is a maximum of 25 records per institution. Records 1 through 12 define age categories for full-time students and records 13 through 24 contain the same age categories for part-time students. Record 25 is the grand total. Each record contains enrollment data by sex of student for undergraduate, first-professional and graduates student levels. The records for EF2 institutions are similar to EF1 institutions but contains enrollment data for undergraduate students only. There are no CN records in this data file. *This file is available in odd-numbered years only.*

The file structure for part D (clarifying questions file:EF97\_D, has one record per institution for both the EF1 and EF2 surveys. (see record layout).

A more detail description of the line number by survey form for parts A and B is provided in appendix B. For a more detailed description of each variable please refer to the attached record layout.

## Data Analysis: Subsetting the Records

There are obviously many ways of subsetting the file. The point here is to demonstrate the kind of logic one must use to obtain the proper records for analysis. Just remember that the ordering of the IF tests (or variable selection for non-SAS users) can dramatically affect processing time. The rule of thumb is to eliminate all inapplicable records at the outset for maximum efficiency in processing, so consider whether sorting variables from most general to most specific reduces the amount of time it takes to subset the file, and structure the selection of cases accordingly.

In order to accurately subset the file, a programmer must determine the type and control of institutions included in the analysis and the record positions of desired variables. For example, suppose the following question is asked: "What is the racial/ethnic distribution of all full-time undergraduate students in Title IV eligible, degree-granting public 4-year institutions in the 50 states and the District of Columbia?"

Using SAS statements as an example, selecting the proper records could be performed as follows:

```
DATA EF97_IC;
INFILE 'F:\IPED9798\FINAL\EF97_IC.DAT' LRECL=136 RECFM=V MISSEVER;
INPUT
@001 UNITID          6.
@079 FIPS            2.
@097 SECTOR          2.
@130 OPEIND          1.
@131 HDEGOFFR        $3. ;      Select the variable from the IC file module.
IF SECTOR=1 AND OPEIND IN (1 2) and HDEGOFFR > '00' AND FIPS LE 56;
    SECTOR=1 selects 4-year public institutions.
    OPEIND = 1 OR 2, selects Title IV eligible institutions,
    HDEGOFFR > '00' select degree-granting institutions,
    FIPS le 56 selects only institutions in the U.S. and D.C..
```

```
DATA EF97_ANR;
INFILE 'F:\IPED9798\FINAL\EF97_ANR.DAT' LS = 136 RECFM=V MISSEVER;
INPUT @001 UNITID  6.
@007 PART $1.
@008 LINE $2.
@010 LSTUDY  $1.
@011 SECTION  $2.
@013 EFRACE01 6.
@019 EFRACE02 6.
@025 EFRACE03 6.
@031 EFRACE04 6.
@037 EFRACE05 6.
@043 EFRACE06 6.
@049 EFRACE07 6. Define race ethnic record.
```

@055 EFRACE08 6.  
@061 EFRACE09 6.  
@067 EFRACE10 6.  
@073 EFRACE11 6.  
@079 EFRACE12 6.  
@085 EFRACE13 6.  
@091 EFRACE14 6.  
@097 EFRACE15 6.  
@103 EFRACE16 6.  
@109 BALMEN 6.  
@115 BALWOM 6.;

IF line = 8; select full time undergraduate line.  
Now merge the files.

```
Data ef;  
Merge ef97_ic(in=Inic)  
      ef97_anr(in=ina);  
by UNITID
```

if inic; select only those enrollments for the institutions selected in the first step.

Similar analysis can be done using several other programming languages such as SPSS. Data base languages such as Microsoft's ACCESS can also be used.

### **Cautionary Notes**

The definitions and instructions for compiling IPEDS data have been designed to minimize problems encountered when making institutional comparisons. However, postsecondary institutions differ widely among themselves. In addition, imputed data are on file for institutions that were partial or total nonrespondents to the survey (RSTATUS=4 codes for total nonrespondents, record position 89. Caution should always be exercised when comparing institutions for which data have been imputed since these data are intended for computing national totals and are not intended to be an accurate portrayal of an institution's data. Thus, use of imputed data may result in misleading comparisons.

One should also use caution when making year to year enrollment comparisons, by state. In some cases, state enrollment counts may vary from year to year as a result of imputation rather than a result of changes in the reported enrollment data. To prevent misinterpretation, always check the RSTATUS codes of individual institutions to determine if a large proportion of data were imputed when considering year to year changes, by state.

In addition, caution is urged in using the information on Part D of the form concerning remedial enrollment. This question is geared towards those students who are enrolled exclusively in remedial courses - not students enrolled in one or more such courses.

Another note of caution concerns the use of form type (EF1, EF2, CN) vs. institutional sector. Forms were mailed to institutions based on information provided on the prior year's Institutional Characteristics (IC) survey; i.e., control and highest level of offering, which when combined, determine an institution's sector. For example, an accredited public 2-but-less-than-4-year school (sector 4) should receive an EF2 survey form. In some cases, these schools elected to return an EF1 form, either because they now offer a 4-year award, or because they may have upper division or postbaccalaureate students taking courses there. When schools returned forms that were inconsistent with the sector in which they were identified, NCES attempted to determine their proper sector and either the school's sector was adjusted, or the data returned were adjusted to conform with the proper survey form.

While IPEDS encourages institutions to report data for each branch campus separately, some institutions choose to report all their data with their main campus "i.e., parent institution". Institutions that report their data with another institution are known as child institutions and are identified on the file by the variable PARCHILD=2. Institutions that report the combined data (parent institutions) are identified by PARCHILD=1. Parent and child institutions have a data field called UNITIDX, which contains the UNITID of the parent institution record. This variable allows users to link the child record with the parent record.