Standard Operating Procedures
Post-Secondary Employment Outcomes (PSEO)
Demographic File

Last Updated: February 3, 2025

I.   Purpose

This operating procedure specifies the standard layout of the demographic file that partners will supply for use in processing and producing the Post-Secondary Employment Outcomes (PSEO). However, in contrast to the Degree File SOP, some parts of this file are provided as a guideline for institutions as opposed to a strict file specification.

II.   Notes

The most important variable of this file is *residency*. We intend to use residency to provide tabulations of employment flows for students who began in or out of state. For this reason, residency ought to be based on *geographic residency*, not residency for tuition purposes. We also ask for specific values for residency. *For students attending a private institution, it is unlikely the institution collects residency, so all values should be "N" (Not collected), rather than "0" (Out-of-state).*

In addition to residency, there are other variables which will only be available for certain students. We expect missing data in these fields (e.g., SAT scores for a doctoral student). We explicitly mention when missingness is expected. Missing SSNs are not an issue, so transfer all records, regardless of availability of SSN for that observation.

III.   Filename Convention

Use SYSTEM_dmg_STARTYR_ENDYR.txt where SYSTEM is the agreed upon abbreviation for your school system; STARTYR is the first year of the data you are sending; and ENDYR is the last year of the data you are sending. For example, utsys_dmg_2001_2015.txt is the file name for the University of Texas System's input file with data from 2001 to 2015. Data for a single academic year should still include two years. For example, utsys_dmg_2001_2002.txt for academic year 2001-2002.

IV.   Observation Level

An observation in this dataset is a unique individual-institution-degree level. Residency ought to be determined based on first year of entry into that degree level, since there are transitions, especially during graduate school.

V.   Valid Years

We request as many years of historical data as you can send, starting with 2001. If you are unable to go back to 2001, that is fine. We will not process any years before 2001 because of coverage issues with our wage data before that calendar year. If you are sending new years of data, please do not resend data for years from a prior submission, unless explicitly asked to.

VI.   Contact

Andrew Foote	(202) 384-4784	andrew.foote@census.gov
Cody Orr	(202) 961-9780	cody.orr@census.gov

LEHD PSEO Demographic File Format

Label	Data Element	Len	Type	Data Specification
ssn	Social Security Number (SSN)	9	Character	A 9-digit code indicating each student's Social Security Number. *Do not include hyphens. Leave blank if unknown.*
name_first	First Name	50L	String	The first name of the reference student, if known. *Left justify, blank fill.*
name_last	Last Name	50L	String	The last name (surname) of the reference student, if known. *Left justify, blank fill.*
dob	Date of Birth	8	String YYYYMMDD	The date of birth of the reference student, if known, in YYYYMMDD. *Do not include hyphens or other special characters. Leave blank if unknown.*
deglevl_code	Degree/Award Level	2R	String	A 2-digit code indicating the reference student's degree/award level. For acceptable values, see link. *Right justify, zero fill.*
opeid	Office of Postsecondary Education Identification (OPEID) Number	8R	String	An 8-digit identification number indicating the reference student's institution. For acceptable values, see link. *Right justify, zero fill.*
sex	Sex of student			Sex as reported to IPEDS. Can be codes from your own system, but you will need to provide a codebook of the values if not self-explanatory.
race_ethnicity	Race/ethnicity of Student			Race/ethnicity combined as reported to IPEDS. Can be codes from your own system, but you will need to provide a codebook of the values if not self-explanatory. The current IPEDS categories are:

				"American Indian or Alaska Native", "Asian", "Black or African American", "Hispanic or Latino", "Native Hawaiian or Other Pacific Islander", "White", "Two or more race", "Race/ethnicity unknown", "U.S. Nonresident".
residency	Residency status	1	Character	Geographic residency upon entry into institution and degree program. Valid values: 1 = In-state resident upon entry into degree, 0 = Not in-state resident, N = Not collected. *If institution is a private school, it is unlikely they collect this information, and all values should be N, rather than 0. For more instructions regarding residency, see section II above.*
hs_attended	High school attended			ETS or CEEB code. *Can be missing.*
sat_act_score	SAT/ACT Scores			Preferably separated by field (math/English) and composite. You may use any number of variable names to denote the individual fields as long as they are self-explanatory. *Can be missing.*