

NAWS Public Access Data: Fiscal Years 1989 – 2009

Background

This version of the National Agricultural Workers Survey (NAWS) Public Access Data (NAWSPAD) contains information from 52,479 in-person interviews with hired crop farm workers. The interviews were conducted in 467 counties and 40 states between October 1, 1988 and September 30, 2009. The data set includes 220 questionnaire variables and 100 created variables. This document describes the data and provides analysis tips.

The data are currently available in a zipped SAS file. Summary tables of the survey's key variables will be posted to the NAWS Web page this year. It is highly recommended that the questionnaire and codebook be downloaded and examined prior to analyzing data. Questionnaire variable names in the codebook match the cycle 64 questionnaire.

The NAWSPAD was constructed under two guiding principles: 1) to protect the privacy rights of the farm worker respondents and the agricultural employers at whose establishments the workers were interviewed; and 2) to provide a data set that allows in-depth analysis of the demographic and employment characteristics of hired crop workers at both the national and regional level.

Confidential Data Safeguards

The names and addresses of employers and respondents do not appear in either the full NAWS data set or the NAWSPAD. Additional safeguards have been incorporated in the NAWSPAD: the variables 'crop' and 'task' have been collapsed into five and six categories, respectively, the 12 sampling regions have been collapsed into six, and the season of interview variable has been removed.

Sampling Regions

Workers are sampled from 12 regions, which are aggregated from 17 U.S. Department of Agriculture-designated regions. The sampling methodology is discussed at <http://www.doleta.gov/agworker/statmethods.cfm>. As noted above, the 12 regions have been collapsed into six in the NAWSPAD. Table 1 lists the regions and the states each region represents; Table 2 shows the number of interviews by fiscal year (FY) and region.

Table 1. NAWSPAD Regions and States

Region No.	Region Name	States Represented
1	EAST	North Carolina, Virginia, Kentucky, Tennessee, West Virginia, Connecticut, Maine, Massachusetts, New Hampshire, New York, Rhode Island, Vermont, Delaware, Maryland, New Jersey, Pennsylvania
2	SOUTHEAST	Arkansas, Louisiana, Mississippi, Alabama, Georgia, South Carolina, Florida
3	MIDWEST	Illinois, Indiana, Ohio, Iowa, Missouri, Kansas, Nebraska, North Dakota, South Dakota, Michigan, Minnesota, Wisconsin
4	SOUTHWEST	Arizona, New Mexico, Oklahoma, Texas
5	NORTHWEST	Idaho, Montana, Wyoming, Colorado, Nevada, Utah, Oregon, Washington
6	CALIFORNIA	California

Table 2. Number of Interviews by Fiscal Year and Region, 1989-2009

FY	Region 1 East	Region 2 Southeast	Region 3 Midwest	Region 4 Southwest	Region 5 Northwest	Region 6 California	Total
1989	311	504	322	269	324	796	2526
1990	149	571	58	300	230	810	2118
1991	149	731	88	321	279	1032	2600
1992	147	552	87	235	201	674	1896
1993	423	351	442	248	265	664	2393
1994	478	320	454	221	290	611	2374
1995	347	306	411	237	354	708	2363
1996	264	371	306	242	277	591	2051
1997	338	397	257	242	264	602	2100
1998	319	428	262	236	262	592	2099
1999	569	487	546	239	489	1282	3612
2000	383	603	440	223	528	1409	3586
2001	414	548	393	177	417	1162	3111
2002	445	585	478	224	431	1198	3361
2003	464	616	499	266	477	1263	3585
2004	413	519	455	171	407	1081	3046
2005	314	337	374	118	314	770	2227
2006	176	300	196	116	196	535	1519
2007	197	294	176	100	193	551	1511
2008	316	407	211	142	265	841	2182
2009	275	392	232	162	337	821	2219
Total	6891	9619	6687	4489	6800	17993	52,479

Interview Cycles and Analysis Tips

Interviews are conducted three times per year to account for the seasonality of agricultural production and employment. The cycles begin in October, February, and July. Due to the timing of the cycles, the data are analyzable by Federal fiscal year (October 1 to September 30) rather than calendar year. The data are not analyzable by interview cycle.

Sampling and post-sampling weights are used in the NAWS to adjust the relative value of each interview so that population estimates may be obtained from the sample. The composite weight variable, PWTYCRD, is included in the NAWSPAD and it should be used in all analyses to generate unbiased percentages and means. For example, to obtain the percentage of male and female crop workers in Federal fiscal years 2001-2002, the following SAS statements would be used:

```
proc freq data= (name of data set);  
table GENDER; where 2001 le FY le 2002;  
weight PWTYCRD;  
run;
```

For a fuller explanation of the weights, please consult pages 6-13 of the statistical methods document: <http://www.doleta.gov/agworker/statmethods.cfm> .

Data Limitations

Except for California, the data are not available at the state-level. Due to low annual sample sizes, at least four years should be combined for all regional-level analyses. When analyzing data at the national level, at least two years of data should be combined.

Due to the survey's complex sampling design, a special analysis procedure is needed to generate survey-design-corrected standard errors. The procedure used in SAS is discussed on pages 17-18 of the statistical methods document. As the procedure requires the variable for the primary sampling units, which, for confidentiality safeguards, is not included in the NAWSPAD, researchers who need to report design-corrected standard errors are encouraged to contact Mr. Daniel Carroll via phone (202) 693-2795 or e-mail (carroll.daniel.j@dol.gov) to discuss their information needs.

The health condition variables provided in the NAWAPAD represent diagnoses over the life of the crop worker and may not represent existing conditions for the farm worker at the time of interview. Work-relatedness was not assessed for these health conditions.