

**MEPS HC-025:  
1998 Jobs File**

**October 2001**

**Agency for Healthcare Research and Quality  
Center for Cost and Financing Studies**

## Table of Contents

A. Data Use Agreement .....	A-1
B. Background .....	B-1
1.0 Household Component .....	B-1
2.0 Medical Provider Component .....	B-2
3.0 Insurance Component .....	B-3
4.0 Survey Management .....	B-3
C. Technical and Programming Information .....	C-1
1.0 General Information .....	C-1
2.0 Data File Description .....	C-1
2.1 Codebook Structure .....	C-4
2.2 Reserved Codes .....	C-4
2.3 Codebook Format .....	C-4
2.4 Variable Source and Naming Conventions .....	C-4
D. Variable-Source Crosswalk .....	D-1
Appendix 1: Sample SAS Program .....	A1-1

## **A. Data Use Agreement**

Individual identifiers have been removed from the microdata contained in the files in this release. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and/or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which they were supplied; any effort to determine the identity of any reported cases is prohibited by law.

Therefore in accordance with the above referenced Federal statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis.
2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director, Office of Management, AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity.
3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey.

By using these data you signify your agreement to comply with the above-stated statutorily based requirements, with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates 18 U.S.C. 1001 and is punishable by a fine of up to \$10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

## **B. Background**

This documentation describes one in a series of public use files from the Medical Expenditure Panel Survey (MEPS). The survey provides an extensive data set on the use of health services and health care in the United States.

MEPS is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ) (formerly the Agency for Health Care Policy and Research (AHCPR)) and the National Center for Health Statistics (NCHS).

MEPS comprises three component surveys: the Household Component (HC), the Medical Provider Component (MPC), and the Insurance Component (IC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES, also known as NMES-1) was conducted in 1977. The National Medical Expenditure Survey (NMES-2) was conducted in 1987. Beginning in 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To accommodate these goals, new MEPS design features include linkage with the National Health Interview Survey (NHIS), from which the sampling frame for the MEPS HC is drawn, and continuous longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents, medical providers, employers, and insurance providers.

### **1.0 Household Component**

The MEPS HC, a nationally representative survey of the U.S. civilian noninstitutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2 ½ -year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for two calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS HC is drawn from respondents to NHIS, conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian noninstitutionalized population, with oversampling of Hispanics and blacks.

## **2.0 Medical Provider Component**

The MEPS MPC supplements and validates information on medical care events reported in the MEPS HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all hospitals, hospital physicians, home health agencies, and pharmacies reported in the HC. Also included in the MPC are all office-based physicians who:

- were identified by the household respondent as providing care for HC respondents receiving Medicaid.
- were selected through a 75-percent sample of HC households receiving care through an HMO (health maintenance organization) or managed care plan.
- were selected through a 25-percent sample of the remaining HC households.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9-CM (Health Care Financing Administration, 1980) and DSM-IV (American Psychiatric Association, 1994).
- Physician procedure codes classified by CPT-4 (Common Procedure Terminology, Version 4).
- Inpatient stay codes classified by DRGs (diagnosis-related groups).
- Prescriptions coded by national drug code (NDC), medication name, strength, and quantity dispensed.
- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials. In some instances, providers sent medical and billing records, which were abstracted into the survey instruments.

### **3.0 Insurance Component**

The MEPS IC collects data on health insurance plans obtained through employers, unions, and other sources of private health insurance. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics.

Establishments participating in the MEPS IC are selected through four sampling frames:

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.
- A Bureau of the Census list frame of private-sector business establishments.
- The Census of Governments from Bureau of the Census.
- An Internal Revenue Service list of the self-employed.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and insurance providers) are linked back to data provided by the MEPS HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance.

The MEPS IC is an annual survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone followup for nonrespondents.

### **4.0 Survey Management**

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files.

Printed documents and CD-ROMs are available through the AHRQ Publications Clearinghouse.  
Write or call:

AHRQ Publications Clearinghouse  
Attn: (publication number)  
P.O. Box 8547  
Silver Spring, MD 20907  
800/358-9295  
410/381-3150 (callers outside the United States only)  
888/586-6340 (toll-free TDD service; hearing impaired only)

Be sure to specify the AHRQ number of the document or CD-ROM you are requesting. Selected electronic files are available from the Internet on the MEPS web site:  
<<http://www.meps.ahrq.gov/>>.

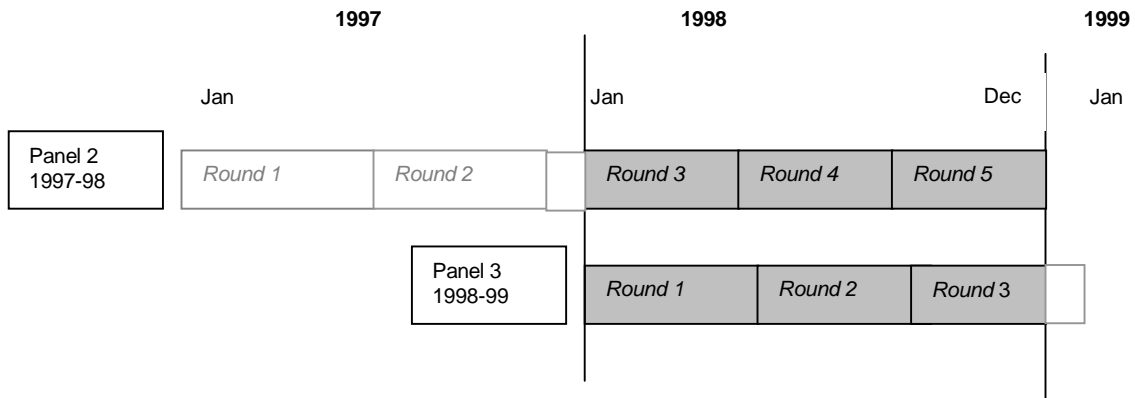
Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Healthcare Research and Quality.

## C. Technical and Programming Information

Section C of this document offers a brief overview of the data provided in MEPS public use release HC-025, as well as the content and structure of the codebook, reserved code values and variable name conventions. It is followed by Section D containing the Variable-Source Crosswalk, and Appendix 1 containing sample SAS program code. A copy of the survey instrument used to collect the information on this file is available on the MEPS web site : <http://www.meps.ahrq.gov>.

### 1.0 General Information

The Jobs file provided in this release, MEPS HC-025, contains job level information collected in Rounds 3-5 for the second Panel and Rounds 1-3 for the third Panel of the 1998 Medical Expenditure Panel Survey (MEPS), as illustrated below.



It is being released as a research file and has not undergone the standard quality control procedures usually performed on MEPS public use data files. The file includes a total of 41,406 records, with each record representing a unique job for a person by round. For this file, only information for jobs starting on or before 12/31/1998 is presented. Information on jobs starting in 1999 will be contained in the 1999 MEPS Jobs file release.

### 2.0 Data File Description

The unique record identifier is the variable JOBSIDX which is comprised of a person identifier (DUID + PID), a round identifier (RN), and a job number (JOBSN). A panel indicator (PANEL) is included on the files to distinguish Round 3 jobs held by Panel 2 persons from Round 3 jobs held by those in Panel 3.

Each job record contains original variables related to the employment section of the 1998 MEPS household survey. All persons aged 16 and older in the MEPS can report any number of jobs held within a round. Persons who held more than one job at the round's interview date (a current job) were asked to identify the main job. This job was classified as the current main job and all other simultaneously held jobs were classified as miscellaneous. The MEPS also obtained some



information on prior jobs held in the reference period and, for those persons not currently working and having no job in the reference period, some information on the last job the person held. Additionally, for those persons age 55 and older who indicated that they retired from a job, the MEPS obtained some job-level data on the retirement job. The variable SUBTYPE indicates the type of job record – current main, current miscellaneous, former main, former miscellaneous, last job outside reference period, or retirement job. It is important to note that the retirement job classification in the variable SUBTYPE is independent of the retirement response in the variable relating to the question why a person left a previous job (YLEFT). The MEPS asked for detailed information about any job classified as current main or current miscellaneous and basic information about any other job types. Refer to the questionnaire to see which information was asked for each job type.

The MEPS used dependent interviewing in Rounds 3, 4, and 5 for Panel 2 and in Rounds 2 and 3 for Panel 3. In these rounds, the MEPS asked persons who held current main and current miscellaneous jobs at the previous round interview date whether they were still working at these jobs (see section RJ in the employment section). If a person still held a current main job from the previous round, the MEPS asked whether the job was still the main job. Most persons reported that they still worked at the same job and it was still their main job. If job status remained the same, the MEPS asked only a subset of the employment questions. Because the MEPS asked only this subset of questions if job status for a person did not change in later rounds, many job level variables on the subsequent round job records are coded as inapplicable; the complete information is on the jobs record for the job in a previous round. Thus it is important to determine whether a job in a subsequent round continues from the previous round when working with the job records. In rounds where this applies, the variables STILLAT (for jobs that were current main in the previous round) and STILLWRK (for jobs that were current miscellaneous in the previous round) indicate whether a person still holds the job at the subsequent round interview date. The variable SUBTYPE on the subsequent round job record indicates whether the job is main or miscellaneous in that subsequent round. Note that if a Panel 2 Round 3-5 job is continued from a Round 1 or 2 job, much of the information will be contained in the 1997 MEPS Jobs file (HC-019). Use that file to obtain the desired job characteristics. A sample SAS program showing how to do this is provided in Appendix 1. Variables that relate only to the review of a job reported in a previous round (Y\_CHANGE, OFFTAKEI, NOWTAKEI, WHY\_LEFT, STILLAT, STILLWRK, DIFFWAGE, WHY\_DIFF, WORKSTAT, and SHFTCHNG) were not asked in Round 1 and these variables are coded as inapplicable.

There are two variables on the file which, while skipped over as inapplicable during the actual interview process in rounds subsequent to the one in which the job was initially reported, have their original value carried forward on the database from round to round. The variables are MORELOC and NUMEMPS. This treatment is due to special internal processing needs. Unlike the situation for many other variables, as explained above, values other than inapplicable for MORELOC and NUMEMPS do not necessarily mean that a job is newly reported, as they may be carried forward from an earlier round.

In the review section (RJ), the MEPS attempted to obtain information regarding changes in wages for the same job from round to round; however in many cases, but not all, wage changes of less

than \$.50 per hour on the same job were not recorded. This caution applies only to individual jobs that did not change. For every new job reported for a person, the MEPS attempted to obtain up-to-date wage information.

For reasons of confidentiality, the hourly wage variable (HRLYWAGE) was topcoded. A value of -10 indicates that the hourly wage rate was greater than or equal to \$76.00. Additionally, the value of any other earnings on a job (BONSAMT, COMMAMT, TIPSAMT, DAYWAGE, WKLYAMT, GROSSPAY, MAKEAMT, OTHRWAGE) with a calculated hourly wage greater than or equal to \$76.00 was also capped. To calculate the hourly wage for these other types of compensation, the number of hours per week worked and in some cases the number of weeks worked were used in conjunction with the various amounts. (These calculated hourly wages do not appear on the file.) Compensation variables were not reconciled with income data collected elsewhere in the MEPS. Additionally, for confidentiality, establishment size variables (NUMEMPS - jobs held by wage earners, ESTMATE1 - estimate of establishment size for jobs held by wage earners, and TOTLEMP - self-employed jobs) were topcoded as -10 for sizes greater than or equal to 10,000.

Due to many skip patterns in the instrument, it is recommended that users of these data become familiar with the employment section in the MEPS questionnaire. To aid users, a crosswalk between variables and MEPS questionnaire numbers is provided in this release. The following examples of variables involved in skip patterns are presented to be illustrative; these examples do not represent the full range of variables affected by questionnaire skip patterns. In one example of a skip pattern, the MEPS did not obtain job-related benefits such as vacation, sick leave and pension information for self-employed jobs, so those variables are coded as inapplicable (-1) for those types of jobs. Nor did the MEPS attempt to obtain wage, salary, and information regarding whether the job was in the private sector, federal or local government (TYPEEMPL) for the self-employed. So again, due to the skip pattern, TYPEEMPL is coded as inapplicable for self-employed jobs. Conversely, the questions relating to business organization type (BUSINC, PROPRIET) were asked only of the self-employed, so the skip pattern results in those variables being coded inapplicable for jobs performed by wage earners.

It is important to note that the establishment size variable for the self-employed is TOTLEMP while the establishment size for wage earners can be found in NUMEMPS and ESTMATE1. The variable ESTMATE1 is derived from a question that allowed wage earners who did not know the actual establishment size (NUMEMPS) to choose from a number of size ranges. Industry codes (INDTCODX) and occupation codes (OCCPCODX) were coded by professional coders from verbatim responses.

This file does not include any weights. To make person level estimates, link to any of the 1998 files and use the person level weight for the appropriate panel. The link should be made through the variable DUPERSID. Note that not all persons in the MEPS have positive weights.

## 2.1 Codebook Structure

For each variable on the file, an unweighted frequency is provided in the codebook, file H25CB.PDF, on the MEPS web site: <<http://www.meps.ahrq.gov>>.

## 2.2 Reserved Codes

The following reserved code values are used:

Value	Definition
-1 INAPPLICABLE	Question was not asked due to skip pattern.
-7 REFUSED	Question was asked and respondent refused to answer question.
-8 DK	Question was asked and respondent did not know answer.
-9 NOT ASCERTAINED	Interviewer did not record the data.
-10 TOP-CODED VALUE	Value was top-coded for confidentiality, as described above.

## 2.3 Codebook Format

The codebook describes an ASCII data set (although the data are also being provided in a SAS transport file). The following codebook items are provided for each variable:

Identifier	Description
Name	Variable name (maximum of 8 characters)
Description	Variable descriptor (maximum of 40 characters)
Format	Number of bytes
Type	Type of data: numeric (indicated by NUM) or character (indicated by CHAR)
Start	Beginning column position of variable in record
End	Ending column position of variable in record

## 2.4 Variable Source and Naming Conventions

In general, variable names reflect the content of the variable, with an 8-character limitation. Variables contained on this file were derived from the questionnaire itself or from the CAPI. The source of each variable is identified in [Section D, “Variable - Source Crosswalk”](#). Sources for each variable are indicated in one of two ways:

- (1) variables which are derived from CAPI or assigned in sampling are so indicated as “CAPI derived” or “Assigned in sampling,” respectively;
- (2) variables which come from one or more specific questions have those questionnaire sections and question numbers indicated in the “Source” column.

**D. Variable-Source Crosswalk**

**FOR MEPS PUBLIC USE RELEASE HC-025**

**SURVEY ADMINISTRATION VARIABLES - PUBLIC USE**

<b>VARIABLE</b>	<b>LABEL</b>	<b>SOURCE</b>
JOBSIDX	JOB ID Number	CAPI Derived
DUPERSID	Sample Person ID (DUID+PID)	Assigned in Sampling
PANEL	Panel	Assigned in Sampling
DUID	Dwelling Unit ID	Assigned in Sampling
PID	Person Number	Assigned in Sampling
RN	Round	CAPI Derived
JOBSN	JOB Number	CAPI Derived

**EMPLOYMENT VARIABLES - PUBLIC USE**

<b>VARIABLE</b>	<b>LABEL</b>	<b>SOURCE</b>
JOBTTYPE	Self-Employed or Work for Someone Else	EM05, EM11, EM18, EM27, EM40, EM53, EM70, EM82
JSTRTM	Job Start Date B Month	EM10OV1-2, EM16OV1-2, EM25OV1-2, EM34OV1-2, EM47OV1, EM60OV1-2
JSTRTD	Job Start Date B Day	EM10OV1, EM16OV1, EM25OV1, EM34OV1, EM47OV1, EM60OV1
JSTRTY	Job Start Date B Year	EM10, EM16, EM25, EM34, EM47, EM60
JSTOPM	Job Stop Date - Month	EM35OV1-2, EM48OV1-2, EM61OV1-2, EM66OV1-2, EM89OV1-2, RJ09
JSTOPD	Job Stop Date B Day	EMJ35OV1, EM48OV1, EM61OV1, EM89OV1, EM66OV1, RJ09
JSTOPY	Job Stop Date B Year	EM35, EM48, EM61, EM66, EM89, RJ09
RETIRJOB	Person Retired From This Job	EM80
SUBTYPE	Job Sub-Type	EM08, EM14, RJ01, RJ01A
MAIN_JOB	Still Main Job or Business	RJ01A
DIFFWAGE	Any Change in Wage Amount	RJ02

<b>VARIABLE</b>	<b>LABEL</b>	<b>SOURCE</b>
WHY_DIFF	Why Wages Changed	RJ03
WORKSTAT	Full or Part Time	RJ04
Y_CHANGE	Why Change in Full/Part Time Status	RJ05
STILLWRK	Still Work at Estab/Misc job	RJ01A, RJ06
OFFTAKEI	Offered Insurance and Now Take	RJ07
NOWTAKEI	Now Offered and Take Insurance	RJ08, RJ08A
WHY_LEFT	Reason Why Not at Job Now	RJ10
NUMEMPS	Employees Per Week	EM91
ESTIMATE1	Total Employees	EM92
MORELOC	More Than One Location	EM93
BUSINC	Business Incorporated	EM94
PROPRIET	Proprietorship or Partnership	EM95
TYPEEMPL	Employee Type	EM96
YLEFT	No Job Reason	EM101
YNOBUSN	Why No Business	EM102
RECALL	Recall Within 30 Days	EM103
HRSPRWK	Number of Hours Per Week	EM104, EM111
HRS35WK	Work 35 Hours	EM105
APXHRDAY	Approximate # of Hours Worked Per Day	EM106
SICKPAY	Does Person Have Paid Sick Leave	EM107
PAYDRVST	Is There Paid Sick Leave for Dr's Visits	EM108
PAYVACTN	Does Person Get Paid Vacation	EM109
RETIRPLN	Does Person Have Pension/Retirement Plan	EM110
WKLYAMT	Usual Weekly Gross Income	EM112
EMPLINS	Have Health Insurance Through This Job	EM113
OFFRDINS	Offered Insurance But Chose Not to Take	EM114
DIFFPLNS	Choice of Different Health Ins Plans	EM115
INUNION	Belong to Labor Union	EM116
PROVDINS	Who Provides Health Insurance	EM117

<b>VARIABLE</b>	<b>LABEL</b>	<b>SOURCE</b>
HHMEMBER	Any Other HH Member Work at This Business	EM122
NUMFMEMB	How Many HH Members Work There	EM123
TOTLEMP	Total Number of Employees at That Business	EM124, RJ08B
SALARIED	Is Person Salaried, Paid by the Hour, etc.	EW01
HOWPAID	How Is Person Paid	EW02
DAYWAGE	Person's Daily Wage Rate	EW03
HRSPRDY	Number of Hours Person Worked in One Day	EW04
MAKEAMT	How Much Money Does Person Make	EW05
PERUNIT	Period for Which Person is Paid	EW050V1
MORE10	Person Makes More or Less than \$10/Hour	EW08, EW14, EW20
MORE15	Person Makes More or Less than \$15/Hour	EW09, EW15, EW21
MOREMINM	Person Makes More or Less than Min. Wage	EW10, EW16, EW22
OVRTIMHR	Overtime Pay Rate Per Hour	EW06
GROSSPAY	Person's Salary Before Takes (Gross)	EW11
GROSSPER	Period in Which Gross Salary was Earned	EW110V1
SALRYWKS	Number of Weeks Per Year Salary is Based	EW12
OTHRTYPE	Type of Overtime Pay	EW19
EARNTIPS	Does Person Earn Tips	EW23_01
EARNBONS	Does Person Earn Bonuses	EW23_02
EARNCOMM	Does Person Earn Commission	EW23_03
OTHRWAGE	Overtime Pay Rate Per Hour	EW19OV1
TIPSUNIT	Period Which Tip Earnings are Based on	EW24AOV1
TIPSAMT	How Much Are Person's Tips	EW24A
BONSUNIT	Period Which Bonuses are Based on	EW24BOV1
BONSAMT	How Much Are Person's Bonuses	EW24B
COMMUNIT	Period Which Commissions are Based on	EW24COV1

<b>VARIABLE</b>	<b>LABEL</b>	<b>SOURCE</b>
COMMAMT	How Much Are Person's Commissions	EW24C
HRLYWAGE	How Much Person Makes/Hr	EW07, EW13, EW18
JBEGHR	Job Begin Hour of Day	EM105A, EM111A
JENDHR	Job End Hour of Day	EM105OV, EM111AOV
JBEGMIN	Job Begin Minutes	EM105A, EM111A
JENDMIN	Job End Minutes	EM105OV, EM111AOV
JBEGAMPM	Job Begin AM/PM	EM105A, EM111A
JENDAMPM	Job End AM/PM	EM105OV, EM111AOV
SHIFTWK	Work Rotating Shifts	EM105B, EM111B
JOBHASHI	Does Person Have Health Ins at This Job	EM17, EM26, EM39, EM52, EM69, EM81
STILLAT	Still Work at Estab/Main job	RJ01, RJ01A
SHFTCHNG	Has a Change in Shift Occurred	RJ05A
HRSALBAS	Hours Salary Based On	EW17
INDTCODX	Industry Code	EM98
OCCPCODX	Occupation Code	EM99, EM100

## **Appendix 1: Sample SAS Program**



NOTE: Copyright (c) 1989-1996 by SAS Institute Inc., Cary, NC, USA.  
 NOTE: SAS (r) Proprietary Software Release 6.12 TS060  
 Licensed to SOCIAL & SCIENTIFIC SYSTEMS INC, Site 0026405003.

WARNING: Your system is scheduled to expire on November 15, 2001, which is 28 days from now. Please contact your installation representative to have your system renewed. The SAS system will no longer function on or after that date.

```

1
2      OPTIONS LS=132 PS=79;
3      *-----
4      *
5      *   Program Name:   SAMPLE.SAS
6      *
7      *   Description:   This job provides an example of how to get job info
8      *                   from previous rounds when STILLAT=1 (indicates still
9      *                   at the current main job) on the job record
10     *
11     *                   In this example, the variable SICKPAY is updated
12     *                   based on information from the current main job in
13     *                   the 1997 JOBS file
14     *
15     *-----;
16
17     LIBNAME INOUT   'T:\AHRQ\CCFS\DB009DVK\DATA\';
NOTE: Libref INOUT was successfully assigned as follows:
      Engine:         V612
      Physical Name:  T:\AHRQ\CCFS\DB009DVK\DATA
18     LIBNAME IN2    'T:\AHRQ\CCFS\HCPUFS\HC019\DATA\';
NOTE: Libref IN2 was successfully assigned as follows:
      Engine:         V612
      Physical Name:  T:\AHRQ\CCFS\HCPUFS\HC019\DATA
19
20     TITLE 'EXAMPLE 1';
21
22     *-----
23     *   select panel 2, round 3 main jobs with STILLAT=1, SICKPAY=-1
24     *   from the 1998 jobs file
25     *-----;
26
27     DATA JOBS98;
WARNING: The BASE Product product with which DATASTEP is associated will expire within 30 days. Please contact your SAS
installation representative to have it renewed.
28     SET INOUT.JOBS3X;
29     IF PANEL = 2 & RN=3 & SUBTYPE=1 & STILLAT=1 & SICKPAY=-1;
30     RUN;

NOTE: The data set WORK.JOBS98 has 4365 observations and 88 variables.
NOTE: The DATA statement used 0.6 seconds.

31
32     PROC PRINT DATA=JOBS98 (OBS=10);
WARNING: The BASE Product product with which PRINT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
33     TITLE4 'SUBSET: PANEL=2, RN=3, SUBTYPE=1, STILLAT=1, SICKPAY=-1';
34     TITLE5 '(FROM 1998 JOBS FILE)'; TITLE6;
35     RUN;

NOTE: The PROCEDURE PRINT printed page 1.
NOTE: The PROCEDURE PRINT used 0.05 seconds.

36
37     *-----
38     *   get corresponding jobs records from the 1997 file
39     *   merge by DUPERSID JOBSN to get jobs from all rounds
40     *-----;
41
42     PROC SORT DATA=JOBS98; BY DUPERSID JOBSN;
WARNING: The BASE Product product with which SORT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
43     RUN;

NOTE: The data set WORK.JOBS98 has 4365 observations and 88 variables.
NOTE: The PROCEDURE SORT used 0.33 seconds.

```

```

44
45     PROC SORT DATA=IN2.JOBBS5X OUT=HC019; BY DUPERSID JOBSN; /* 1997 file */
WARNING: The BASE Product product with which SORT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
46     RUN;

```

NOTE: The data set WORK.HC019 has 59548 observations and 88 variables.  
NOTE: The PROCEDURE SORT used 7.63 seconds.

```

47
48     DATA JOBS97;
WARNING: The BASE Product product with which DATASTEP is associated will expire within 30 days. Please contact your SAS
installation representative to have it renewed.
49     MERGE JOBS98 (IN=A KEEP=DUPERSID JOBSN) HC019(IN=B);
50     BY DUPERSID JOBSN;
51     IF A & B;
52     RUN;

```

NOTE: The data set WORK.JOBS97 has 12678 observations and 88 variables.  
NOTE: The DATA statement used 1.32 seconds.

```

53
54     PROC FREQ;
WARNING: The BASE Product product with which FREQ is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
55     TABLE SUBTYPE*STILLAT*SICKPAY RN/LIST MISSING;
56     TITLE4 'MATCHING RECORDS FROM THE 1997 JOBS file'; TITLE5;
57     RUN;

```

NOTE: The PROCEDURE FREQ printed page 2.  
NOTE: The PROCEDURE FREQ used 0.16 seconds.

```

58
59     PROC PRINT DATA=JOBS97 (OBS=15);
WARNING: The BASE Product product with which PRINT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
60     TITLE4 'SAMPLE DUMP OF 1997 JOBS RECORDS'; TITLE5;
61     RUN;

```

NOTE: The PROCEDURE PRINT printed page 3.  
NOTE: The PROCEDURE PRINT used 0.0 seconds.

```

62
63     *-----
64     * drop 1997 records with STILLAT=1 (indicates person is at the
65     * same job, so job info would be recorded in an earlier round)
66     *
67     * also drop 1997 records if SUBTYPE not equal to 1 (indicates
68     * that the job is not the main job)
69     *-----;
70
71     DATA JOBS97;
WARNING: The BASE Product product with which DATASTEP is associated will expire within 30 days. Please contact your SAS
installation representative to have it renewed.
72     SET JOBS97;
73     IF STILLAT=1 OR NOT(SUBTYPE=1) THEN DELETE;
74     RUN;

```

NOTE: The data set WORK.JOBS97 has 4361 observations and 88 variables.  
NOTE: The DATA statement used 0.33 seconds.

```

75
76     PROC PRINT DATA=JOBS97 (OBS=25);
WARNING: The BASE Product product with which PRINT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
77     TITLE4 'SAMPLE DUMP OF 1997 JOBS RECORDS AFTER DELETING';
78     TITLE5 'RECORDS THAT HAVE STILLAT=1, SUBTYPE NE 1'; TITLE6;
79     RUN;

```

NOTE: The PROCEDURE PRINT printed pages 4-5.  
NOTE: The PROCEDURE PRINT used 0.05 seconds.

```

80
81 *-----
82 * select the 1997 record from the last round, if there is
83 * more than one record for the job
84 *-----;
85
86 PROC SORT DATA=JOBS97; BY DUPERSID JOBSN RN;
WARNING: The BASE Product product with which SORT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
87 RUN;

NOTE: The data set WORK.JOBS97 has 4361 observations and 88 variables.
NOTE: The PROCEDURE SORT used 0.27 seconds.

88
89 DATA JOBS97 CHEK;
WARNING: The BASE Product product with which DATASTEP is associated will expire within 30 days. Please contact your SAS
installation representative to have it renewed.
90 SET JOBS97; BY DUPERSID JOBSN;
91 IF LAST.JOBSN THEN OUTPUT JOBS97;
92 IF NOT(FIRST.JOBSN & LAST.JOBSN) THEN OUTPUT CHEK;
93 RUN;

NOTE: The data set WORK.JOBS97 has 4361 observations and 88 variables.
NOTE: The data set WORK.CHEK has 0 observations and 88 variables.
NOTE: The DATA statement used 0.33 seconds.

94
95 PROC PRINT DATA=CHEK(OBS=20);
WARNING: The BASE Product product with which PRINT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
96 TITLE4'DIAGNOSTIC (MORE THAN ONE MATCH)';TITLE5;
97 RUN;

NOTE: No observations in data set WORK.CHEK.
NOTE: The PROCEDURE PRINT used 0.05 seconds.

98
99 PROC PRINT DATA=JOBS97 (OBS=20);
WARNING: The BASE Product product with which PRINT is associated will expire within 30 days. Please contact your SAS installation
representative to have it renewed.
100 TITLE4 'SELECTED 1997 JOBS TO MATCH TO 1998 JOBS'; TITLE5;
101 RUN;

NOTE: The PROCEDURE PRINT printed pages 6-7.
NOTE: The PROCEDURE PRINT used 0.0 seconds.

102
103 *-----
104 * merge 1997 jobs info on to the 1998 jobs record
105 * by DUPERSID JOBSN
106 *
107 * SICKPAYX: SICKPAY from the 1997 jobs record
108 *-----;
109
110 DATA JOBS98;
WARNING: The BASE Product product with which DATASTEP is associated will expire within 30 days. Please contact your SAS
installation representative to have it renewed.
111 MERGE JOBS98(IN=A)
112 JOBS97(IN=B KEEP=DUPERSID JOBSN SICKPAY RENAME=(SICKPAY=SICKPAYX));
113 BY DUPERSID JOBSN;
114 IF A;
115 RUN;

NOTE: The data set WORK.JOBS98 has 4365 observations and 89 variables.
NOTE: The DATA statement used 0.38 seconds.

116
117 PROC FREQ DATA=JOBS98;
WARNING: The BASE Product product with which FREQ is associated will expire within 30 days. Please contact your SAS installation

```

representative to have it renewed.

```
118     TABLE SICKPAY*SICKPAYX /LIST MISSING;
119     TITLE4 '1998 JOBS RECORDS, AFTER MERGING ON 1997 DATA';
120     TITLE5; TITLE6;
121     RUN;
```

NOTE: The PROCEDURE FREQ printed page 8.

NOTE: The PROCEDURE FREQ used 0.05 seconds.

122

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414



MATCHING RECORDS FROM THE 1997 JOBS file

SUBTYPE	STILLAT	SICKPAY	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	-1	-9	14	0.1	14	0.1
1	-1	-8	114	0.9	128	1.0
1	-1	-7	4	0.0	132	1.0
1	-1	-1	591	4.7	723	5.7
1	-1	1	2236	17.6	2959	23.3
1	-1	2	1402	11.1	4361	34.4
1	1	-1	8082	63.7	12443	98.1
2	-1	-1	29	0.2	12472	98.4
3	-1	-9	1	0.0	12473	98.4
3	-1	-8	2	0.0	12475	98.4
3	-1	-1	8	0.1	12483	98.5
3	-1	1	17	0.1	12500	98.6
3	-1	2	59	0.5	12559	99.1
3	2	-1	3	0.0	12562	99.1
4	-1	-1	8	0.1	12570	99.1
5	-1	-1	104	0.8	12674	100.0
6	-1	-1	4	0.0	12678	100.0

ROUND

RN	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	3956	31.2	3956	31.2
2	4361	34.4	8317	65.6
3	4361	34.4	12678	100.0



SAMPLE DUMP OF 1997 JOBS RECORDS AFTER DELETING RECORDS THAT HAVE STILLAT=1, SUBTYPE NE 1

OB	SS	IB	SD	ON	XS	DU	DI	PI	RN	LE	EM	DD	YY	MM	DD	ES	ST	UB	NI	FW	WA	DI	YR	CK	HL	AL	WA	KA	LE	FP	ES	UT	MA		
1	30001030	1	30001030101	30001	30	1	2	2	-1	-1	1993	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-8	-8
2	30005015	1	30005015101	30005	15	1	2	2	-1	-1	1987	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	100	-1	
3	30005022	1	30005022201	30005	22	2	2	2	9	2	1997	-1	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	28	-1	
4	30008018	1	30008018101	30008	18	1	2	1	-1	-1	1985	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
5	30008025	1	30008025101	30008	25	1	2	2	-1	-1	1974	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	40	-1	
6	30009020	1	30009020101	30009	20	1	2	1	-1	-1	1984	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
7	30011017	2	30011017202	30011	17	2	2	2	8	11	1997	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-8	5	
8	30013018	1	30013018101	30013	18	1	2	1	4	12	1997	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
9	30013025	1	30013025101	30013	25	1	2	2	-1	-1	1993	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	12	-1	
10	30014014	1	30014014101	30014	14	1	2	2	-1	-1	1977	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-8	5
11	30022012	1	30022012101	30022	12	1	2	2	-1	-1	1994	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	25	-1	
12	30023015	1	30023015101	30023	15	1	2	2	4	-8	1997	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	10	-1	
13	30023022	1	30023022101	30023	22	1	2	2	-1	-1	1995	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-8	-10
14	30025028	1	30025028201	30025	28	2	2	2	1	-1	1996	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-8	-8

OB	SS	IB	SD	ON	XS	DU	DI	PI	RN	LE	EM	DD	YY	MM	DD	ES	ST	UB	NI	FW	WA	DI	YR	CK	HL	AL	WA	KA	LE	FP	ES	UT	MA				
1	1	-1	-1	2	-1	-1	-1	50	-1	-1	1	1	1	2	-1	1	-1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1			
2	1	-1	-1	1	-1	-1	-1	50	-1	-1	1	2	1	2	-1	1	-1	1	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1		
3	2	-1	-1	1	-1	-1	-1	40	-1	-1	1	2	1	1	-1	1	-1	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1		
4	-1	2	1	-1	-1	-1	-1	30	-1	-1	-1	-1	-1	-1	2	2	-1	2	-1	2	-1	2	-1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
5	1	-1	-1	1	-1	-1	-1	40	-1	-1	1	2	1	1	-1	1	-1	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
6	-1	1	-1	-1	-1	-1	-1	100	-1	-1	-1	-1	-1	-1	1	-1	1	2	-1	2	-1	2	-1	30	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
7	2	-1	-1	1	-1	-1	-1	44	-1	-1	-8	-1	1	2	-1	2	2	-1	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
8	-1	1	-1	-1	-1	-1	-1	60	-1	-1	-1	-1	-1	-1	2	2	-1	2	-1	2	-1	8	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
9	2	-1	-1	1	-1	-1	-1	38	-1	-1	1	1	1	2	-1	1	-1	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
10	1	-1	-1	1	-1	-1	-1	40	-1	-1	1	2	1	1	-1	1	-1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
11	1	-1	-1	1	-1	-1	-1	40	-1	-1	1	1	1	2	-1	1	-1	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
12	2	-1	-1	1	-1	-1	-1	37	-1	7	2	-1	2	2	-1	2	2	-1	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
13	1	-1	-1	1	-1	-1	-1	40	-1	-1	1	2	1	1	-1	1	-1	1	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
14	-8	-1	-1	1	-1	-1	-1	20	-1	-1	2	-1	2	2	-1	2	-1	8	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

OB	SS	IB	SD	ON	XS	DU	DI	PI	RN	LE	EM	DD	YY	MM	DD	ES	ST	UB	NI	FW	WA	DI	YR	CK	HL	AL	WA	KA	LE	FP	ES	UT	MA			
1	400	4	-1	-1	2	2	2	-1	-1	-1	-1	-1	-1.00	9	7	0	0	AM	PM	2	-1	-1	-1	40	12	4										
2	-1	-1	-1	3	2	2	2	-1	-1	-1	-1	-1	-1	12.05	8	5	30	0	AM	PM	2	-1	-1	-1	-1	5	1									
3	-8	-1	-1	-1	2	2	2	-1	-1	-1	-1	-1	-1	-8.00	8	4	0	0	AM	PM	2	-1	-1	-1	40	11	4									
4	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	4	7	0	0	PM	PM	2	-1	-1	-1	-1	1	10									
5	-7	-1	-1	-1	2	1	2	-1	-1	-1	-1	-1	-1	-1.00	8	4	30	30	AM	PM	2	-1	-1	-1	-1	37	7	4								
6	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	7	7	30	30	AM	PM	2	-1	-1	-1	-1	8	2									
7	-1	-1	-1	3	2	2	2	-1	-1	-1	-1	-1	-1	8.00	7	3	0	0	AM	PM	2	-1	-1	-1	-1	4	6									
8	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	4	2	0	0	AM	PM	2	-1	-1	-1	-1	-1	8	2								
9	-1	-1	-1	1	2	2	2	-1	-1	-1	-1	-1	-1	12.00	7	5	0	0	AM	PM	2	-1	-1	-1	-1	11	8									
10	-1	-1	-1	3	2	1	2	-1	-1	-1	-1	-1	-1	15.00	7	4	0	0	AM	PM	2	-1	-1	-1	-1	-1	4	6								
11	-1	-1	-1	1	2	1	2	-1	-1	-1	91	75	-1	-1	8.00	8	5	0	0	AM	PM	2	-1	-1	-1	-1	11	4								
12	-1	-1	-1	1	1	2	2	-1	2	25	-1	-1	-1	3.25	11	5	0	0	AM	PM	2	-1	-1	-1	-1	6	8									
13	-1	-1	-1	3	2	1	2	-1	-1	6	1200	-1	-1	12.82	7	7	0	0	AM	PM	1	-1	-1	-1	-1	-1	4	6								
14	-1	-1	-1	1	2	2	2	-1	-1	-1	-1	-1	-1	5.00	3	7	0	0	PM	PM	2	-1	-1	-1	-1	-1	6	3								







SELECTED 1997 JOBS TO MATCH TO 1998 JOBS

OBS	DIS	NS	JOBS	DI	PI	IR	EN	LE	EM	TD	TY	MD	YO	BE	EB	WF	OR	TC	IF	AL	NR	WH	NU	ES	
16	30028013	1	30028013101	30028	13	1	2	2	-1	-1	1993	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
17	30032029	1	30032029101	30032	29	1	2	2	-1	-1	1991	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
18	30037014	1	30037014101	30037	14	1	2	2	-1	-1	1993	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
19	30037021	1	30037021101	30037	21	1	2	2	6	-1	1996	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
20	30038010	1	30038010101	30038	10	1	2	1	-1	-1	1966	-1	-1	-1	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
<p>MORSE YOE SHCDVILP RFDINORHNTSAHDAHMPMO</p> <p>OBOPYNR RRXXIYYT KMF FNOMMULOA RAE MMR</p> <p>RURPEY OES SHCDVILP RFDINORHNTSAHDAHMPMO</p> <p>ESRELBCCP3RKRARAYLDPNDMMMLRPWPPEURRM</p> <p>OLIIIMEUAR5DPVCPAIIILIIIBE EIAARAANE EIM</p> <p>BONEPFSLWAAASTLMMNNNSONEMMEIGDMII1NH</p> <p>SCCTLTNLK KY YTNNTSSSNRBPDEDET05MR</p>																									
16	1	-1	-1	-8	-1	-1	-1	40	-1	-1	1	1	1	1	-1	1	-1	2	2	-1	-1	-1	-1	-1	-1
17	1	-1	-1	1	-1	-1	-1	41	-1	-1	2	-1	2	-8	-1	2	1	1	2	-1	-1	-1	-1	-1	-1
18	1	-1	-1	1	-1	-1	-1	40	-1	-1	1	1	1	1	-1	1	-1	2	1	1	-1	-1	-1	-1	-1
19	1	-1	-1	1	-1	-1	-1	20	-1	-1	2	-1	2	2	-1	2	-1	2	-1	-1	-1	-1	-1	-1	-1
20	-1	1	-1	-1	-1	-1	-1	60	-1	-1	-1	-1	-1	-1	2	2	-1	2	-1	2	-1	2	-1	4	-1
<p>GRROATAAEAETITOB</p> <p>ROOLHRRRRRHPIIN</p> <p>SSRRRNNNRSUS</p> <p>SSYTBTBCWUSU</p> <p>PPWPYIOOANANAN</p> <p>AEKPPNMGIMIM</p> <p>YRSSESSEMETTT</p>																									
16	-1	-1	-1	-1	2	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
17	-8	-1	-1	-1	2	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
18	-1	-1	-1	3	2	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
19	-1	-1	-1	1	2	1	2	-1	-1	-1	91	25	-1	-1	6.25	5	9	0	0	PM	PM	1	-1	-1	-1
20	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	9	10	0	0	AM	PM	2	-1	-1	-1

## 1998 JOBS RECORDS, AFTER MERGING ON 1997 DATA

SICKPAY	SICKPAYX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-1	.	4	0.1	4	0.1
-1	-9	14	0.3	18	0.4
-1	-8	114	2.6	132	3.0
-1	-7	4	0.1	136	3.1
-1	-1	591	13.5	727	16.7
-1	1	2236	51.2	2963	67.9
-1	2	1402	32.1	4365	100.0