U.S. Department of Justice
Bureau of Justice Statistics

Technical Standards
for Machine-Readable Data

75318

Bureau of Justice Statistics Reports

National Crime Survey:
Criminal Victimization in the United States (annual):
Summary Findings of 1977-78 Changes in Crime and of Trends Since 1973, NCJ-61368
A Description of Trends from 1973 to 1977, NCJ-59898
1978 (final report), NCJ-66480
1977, NCJ-58725
1976, NCJ-49543
1975, NCJ-44593
1974, NCJ-39467
1973, NCJ-34732
The Cost of Negligence: Losses from Preventable Household Burglaries, NCJ-53527

Intimate Victims: A Study of Violence Among Friends and Relatives, NCJ-62319
Crime and Seasonality, NCJ-64818
Criminal Victimization of New York State Residents, 1974-77, NCJ-66481
Criminal Victimization Surveys in 13 American cities (summary report, 1 vol.), NCJ-18471
Boston, NCJ-34818
Buffalo, NCJ-34820
Cincinnati, NCJ-34819
Houston, NCJ-34821
Miami, NCJ-34822
Milwaukee, NCJ-34823
Minneapolis, NCJ-34824
New Orleans, NCJ-34825
Oakland, NCJ-34826
Pittsburgh, NCJ-34827
San Diego, NCJ-34828
San Francisco, NCJ-34829
Washington, D.C., NCJ-34830
Public Attitudes About Crime (13 vols.):
Boston, NCJ-46235
Buffalo, NCJ-46236
Cincinnati, NCJ-46237
Houston, NCJ-46238
Miami, NCJ-46239
Milwaukee, NCJ-46240
Minneapolis, NCJ-46241
New Orleans, NCJ-46242
Oakland, NCJ-46243
Pittsburgh, NCJ-46244
San Diego, NCJ-46245
San Francisco, NCJ-46246
Washington, D.C., NCJ-46247'
Criminal Victimization Surveys in Chicago, Detroit, Los Angeles, New York, and Philadelphia: A Comparison of 1972 and 1974 Findings, NCJ-36360
Criminal Victimization Surveys in the Nation's Five Largest Cities: National Crime Panel Surveys in Chicago, Detroit, Los Angeles, New York, and Philadelphia, 1972, NCJ-16909
Criminal Victimization Surveys in Eight American Cities: A Comparison of 1971/72 and 1974/75 Findings—National Crime Surveys in Atlanta, Baltimore, Cleveland, Dallas, Denver, Newark, Portland, and St. Louis, NCJ-36361
Crimes and Victims: A Report on the Dayton/San Jose Pilot Survey of Victimization, NCJ-013314
Indicators of Crime and Criminal Justice: Quantitative Studies, NCJ-62349

Applications of the National Crime Survey Victimization and Attitude Data:
Public Opinion About Crime: The Attitudes of Victims and Nonvictims in Selected Cities, NCJ-41336
Local Victim Surveys: A Review of the Issues, NCJ-39973

The Police and Public Opinion: An Analysis of Victimization and Attitude Data from 13 American Cities, NCJ-42018
An Introduction to the National Crime Survey, NCJ-43732
Compensating Victims of Violent Crime: Potential Costs and Coverage of a National Program, NCJ-43387
Crime Against Persons in Urban, Suburban, and Rural Areas: A Comparative Analysis of Victimization Rates, NCJ-53551
Rape Victimization in 26 American Cities, NCJ-55878
Criminal Victimization in Urban Schools, NCJ-56396

National Prisoner Statistics:
Capital Punishment (annual):
1978, NCJ-59897
1979 advance report, NCJ-67705
Prisoners in State and Federal Institutions on December 31:
1978, NCJ-64671
1979 advance report, NCJ-66522
Census of State Correctional Facilities, 1974 advance report, NCJ-25642
Profile of State Prison Inmates: Socio-demographic Findings from the 1974 Survey of Inmates of State Correctional Facilities, NCJ-58257
Census of Prisoners in State Correctional Facilities, 1973, NCJ-34729
Census of Jails and Survey of Jail Inmates, 1978, preliminary report, NCJ-55172
Profile of Inmates of Local Jails: Socio-demographic Findings from the 1978 Survey of Inmates of Local Jails, NCJ-65412
The Nation's Jails: A report on the census of jails from the 1972 Survey of Inmates of Local Jails, NCJ-19067
Survey of Inmates of Local Jails, 1972, advance report, NCJ-13313

Uniform Parole Reports:
Parole in the United States (annual):
1978, NCJ-58722
1976 and 1977, NCJ-49702
Characteristics of the Parole Population, 1978, NCJ-66479
A National Survey of Parole-Related Legislation Enacted During the 1979 Legislative Session, NCJ-64218

Children in Custody: Juvenile Detention and Correctional Facility Census
1977 advance report:
Census of Public Juvenile Facilities, NCJ-60967
Census of Private Juvenile Facilities, NCJ-60968
1975 (final report), NCJ-58139
1974, NCJ-57946
1973, NCJ-44777
1971, NCJ-13403

Myths and Realities About Crime: A Nontechnical Presentation of Selected Information from the National Prisoner Statistics Program and the National Crime Survey, NCJ-46249

State and Local Probation and Parole Systems, NCJ-41335
State and Local Prosecution and Civil Attorney Systems, NCJ-41334
National Survey of Court Organization:
1977 Supplement to State Judicial Systems, NCJ-40022
1975 Supplement to State Judicial Systems, NCJ-29433
1971 (full report), NCJ-11427

State Court Model Statistical Dictionary, NCJ-62320
State Court Caseload Statistics:
The State of the Art, NCJ-46934
Annual Report, 1975, NCJ-51885
Annual Report, 1976, NCJ-56599
A Cross-City Comparison of Felony Case Processing, NCJ-55171

Trends in Expenditure and Employment Data for the Criminal Justice System, 1971-77 (annual), NCJ-57463
Expenditure and Employment Data for the Criminal Justice System (annual)
1978 Summary Report, NCJ-66483
1978 final report, NCJ-66482
1977 final report, NCJ-53206

Dictionary of Criminal Justice Data Terminology: Terms and Definitions Proposed for Interstate and National Data Collection and Exchange, NCJ-36747

Justice Agencies in the U.S.: Summary Report of the National Justice Agency List, NCJ-65560
Criminal Justice Agencies in Region
1: Conn., Maine, Mass., N.H., R.I., Vt., NCJ-17930
2: N.J., N.Y., NCJ-17931
3: Del., D.C., Md., Pa., Va., W.Va., NCJ-17932
4: Ala., Ga., Fla., Ky., Miss., N.C., S.C., Tenn., NCJ-17933
5: Ill., Ind., Mich., Minn., Ohio, Wis., NCJ-17934
6: Ark., La., N.Mex., Okla., Tex., NCJ-17935
7: Iowa, Kans., Mo., Nebr., NCJ-17936
8: Colo., Mont., N.Dak., S.Dak., Utah, Wyo., NCJ-17937
9: Ariz., Calif., Hawaii, Nev., NCJ-15151
10: Alaska, Idaho, Oreg., Wash., NCJ-17938

Utilization of Criminal Justice Statistics Project:
Sourcebook of Criminal Justice Statistics 1979 (annual), NCJ-59679
Public Opinion Regarding Crime, Criminal Justice, and Related Topics, NCJ-17419
New Directions in Processing of Juvenile Offenders: The Denver Model, NCJ-17420
Who Gets Detained? An Empirical Analysis of the Pre-Adjudicatory Detention of Juveniles in Denver, NCJ-17417
Juvenile Dispositions: Social and Legal Factors Related to the Processing of Denver Delinquency Cases, NCJ-17418
Offender-Based Transaction Statistics: New Directions in Data Collection and Reporting, NCJ-29645
Sentencing of California Felony Offenders, NCJ-29646
The Judicial Processing of Assault and Burglary Offenders in Selected California Counties, NCJ-29644
Pre-Adjudicatory Detention in Three Juvenile Courts, NCJ-34730
Delinquency Dispositions: An Empirical Analysis of Processing Decisions in Three Juvenile Courts, NCJ-34734
The Patterns and Distribution of Assault Incident Characteristics Among Social Areas, NCJ-40025
Patterns of Robbery Characteristics and Their Occurrence Among Social Areas, NCJ-40026
Crime-Specific Analysis:
The Characteristics of Burglary Incidents, NCJ-42093
An Empirical Examination of Burglary Offender Characteristics, NCJ-43131
An Empirical Examination of Burglary Offenders and Offense Characteristics, NCJ-42476
Sources of National Criminal Justice Statistics: An Annotated Bibliography, NCJ-45006
Federal Criminal Sentencing: Perspectives of Analysis and a Design for Research, NCJ-33683
Variations in Federal Criminal Sentences: A Statistical Assessment at the National Level, NCJ-33684
Federal Sentencing Patterns: A Study of Geographical Variations, NCJ-33685
Predicting Sentences in Federal Courts: The Feasibility of a National Sentencing Policy, NCJ-33686

Technical Standards for Machine-Readable Data
Supplied to the Bureau of Justice Statistics

Report Number NCJ-75318

TABLE OF CONTENTS

Technical Standards for Machine-Readable Data Supplied to the Bureau of Justice Statistics

This standard sets forth technical requirements for data that may be included in the National Criminal Justice Data Archive. The Bureau of Justice Statistics established the archive at the Inter-University Consortium for Political and Social Research, located at the University of Michigan. The archive is designed to promote the analysis of criminal justice data by making inexpensive data tapes available in an easy-to-use form. Machine-readable data files of the bureau's own statistical series as well as other data files that are relevant to national criminal justice issues are available from the archive. To ensure the quality of the data submitted to the archive, the bureau has established these standards.

Any data file that may be submitted for inclusion in the archive must meet these technical standards. Although most of these standards are simply good data processing practice, some of the requirements are specific to the needs and facilities of the archive. Additional information about these standards and the archive or assistance in meeting these standards may be obtained from the bureau.

Another aid to data producers is the bureau's publication, "A Style Manual for Machine-Readable Data Files and Their Documentation," which describes optimal practices and techniques for formatting and documenting machine-readable data files. This document is available as publication NCJ-62766 from the National Criminal Justice Reference Service, Box 6000, Rockville, Maryland, 20850.

Tape recording standards

Data supplied for possible inclusion in the archive should be on 9-track magnetic tape written at a density of 800, 1600, or 6250 bytes (characters) per inch. Since more reliable encoding methods are used at the higher densities, 6250 BPI tapes are preferred to 1600 BPI tapes, which are preferred to 800 BPI tapes. The tape should be labeled with IBM or ANSI volume and file labels. If 9-track equipment is not available, then files may be submitted on 7-track magnetic tape. Either fixed or variable length records are acceptable, and files should be blocked

1

to a length that will provide for efficient handling
of the data. Physical record lengths of from 80 up
to a maximum of 32,767 characters are acceptable.
Experience has shown that physical record lengths of
approximately 16,000 characters provide for the most
efficient processing.[1] This requirement is specific
to the National Criminal Justice Data Archive
computer, which can handle extremely long physical
records. In general, data producers and users should
consult with the technical staff of their computing
facility to determine the facility's preferred size
for physical records. The American National
Standards Institute recommends a maximum physical
record size of 2400 characters, which is considerably
less efficient than the maximum permitted by this
standard. Volume and file-naming conventions should
be agreed on by the bureau and the data supplier
before tapes are written. While it is impractical to
devise a set of file-naming conventions that are uni-
versally applicable, it will often be in the interest
of the bureau and its data suppliers to establish
naming conventions for a particular set of data
files.

Data types

All machine-readable data records supplied for
possible inclusion in the archive will consist en-
tirely of alphanumeric, EBCDIC, or ASCII characters.
(Length fields in binary format that are part of the
IBM variable-length record are not included in this
restriction.)

For the purpose of this standard, variables will
be considered either computational or noncomputation-
al. Noncomputational items contain information such
as names and labels that are never used in any
arithmetic or numerical operation. Noncomputational
variables may contain any EBCDIC or ASCII printing
characters. Noncomputational items will be left

---

[1] Data on magnetic tape are divided into physical and
logical records. A logical record contains data from
a single "unit" (e.g., defendant, gas bill, library
catalog card, etc.). Data processing is often more
efficient if several logical records are written end-
to-end onto the tape to form a single "physical"
record. This process is called "blocking," and a
physical record is sometimes called a "block."

justified and padded to the right with alphabetic
blanks.

Computational items contain numeric information
that is designed to be used in computations.
Computational variables may contain only the char-
acters 0-9, "." "+," and "-." The only exception to
this is where D- or E-format floating-point data are
represented.[2] Wherever possible, data should be
represented in integer format with implicit decimal
places noted in the documentation. Computational
variables should be right justified and padded to the
left with zeros or blanks. If a computational field
is signed, the sign must immediately precede the
left-most numeric character. Computational fields
must contain at least one numeric character. In
particular, computational fields consisting only of a
sign or of blanks are not permissible. Fields con-
taining only a signed zero are not acceptable, since
some computers cannot represent a signed zero.

Missing data. All variables for which there may
be missing data must have an explicit missing data
value or values. In particular, it will never be
assumed that a value of zero or a field consisting
entirely of blanks indicates missing data. Missing
data values must be indicated in the documentation as
a list or a range of values, and may include zero,
but may not include values of -0, a blank field, or a
field consisting only of a sign.

Missing data values must occur in the same field
as the variable to which they refer. If an alternate
value is to be used in place of a missing data value,
the base variable must carry an appropriate missing
data code, while the alternate value will be shown in
a separate variable that has been declared for that

---

[2] Floating-point data are used to represent very
large or very small numbers. The number
2,625,000,000 is written in floating-point notation
as "2.625 E+09." The "E+09" is called the exponent,
and indicates that the decimal point is to be shifted
nine places to the right. D-format floating-point
notation is similar, but uses a "D" rather than an
"E" to mark the exponent. The number
"0.00000000000465" is written in floating-point
notation as "4.65 D-12." the "D" indicates that the
number was computed with double precision arithmetic,
which allows more significant digits to be
represented.

purpose. In no case will the alternate value be
carried in the base variable, with an explanation
code in another variable. The rationale for this
standard is that the meaning of a variable should be
determinable without reference to a second variable.
If the true value of a variable has been suppressed
or modified, then the value of the variable should
indicate such suppression or modification. If an
alternate value is to be offered in such cases, the
appropriate variable in the record can then be read
if the analyst so wishes.[3]

File organization

Variables. Records transmitted for possible
inclusion in the archive should contain no un-
documented or irrelevant fields. The width of a
variable should be sufficient to accommodate the
entire range of variation that may be expected of the
item, but should not be excessive. Conversely,
fields need be no wider than is required to
accommodate the maximum expected value of a variable.
Thus, the variable "number of offenders" for a
multiple victimization incident probably need not be
larger than two characters, and certainly no larger
than three.

In any case, the maximum allowable field width
for an integer variable on tapes submitted to the
bureau is nine characters. Larger numbers should be

---

[3] A brief example may be helpful. Suppose that the
variable "Number of Occupants" is sometimes estimated
rather than counted directly, and that estimated
values are to be flagged. "Number of Occupants"
might be assigned values of 999 for "missing data"
and "998" for "estimated." A second variable,
"Estimated number of occupants" should be used to
supply the estimated value for cases in which "Number
of occupants" is 998. Where "Number of occupants,"
the base variable is not an estimate, "Estimated
number of occupants" should have a value which
denotes "inappropriate." In no case should these
variables be organized so that "Number of occupants"
carries both counted and estimated values, while
"Estimated number of occupants" carries, 0
for observations in which the number of occupants was
counted, and a 1 for observations in which the number
of occupants was estimated.

represented by supplying a scaling factor[4] in the
documentation, or by the use of floating-point
format. A noncomputational variable may be up to
32,767 characters long.

Dates. The Federal Information Processing
Standard (FIPS) format should be used for the
recording of dates in machine-readable form. A date
in FIPS format is expressed as six numeric char-
acters, YYMMDD, where YY is the last two digits of
the year, MM is the ordinal number of the month, and
DD is the day of the month. In this format, February
3, 1983, is written as 830203. Where conciseness is
paramount, dates may be expressed in the alternate
FIPS format of YYDDD, where YY is the last two digits
of the year, and DDD is the ordinal number of the day
in the year. In this format, February 3, 1983, is
written as 83034.

Record type identification. Each record in a
file containing more than one record type, e.g., a
hierarchical file, will carry a variable that
identifies the type of record. The record identifier
must be the first variable in each record, and must
have the same format for all records in the file.
Even where each record in a file has a different
length, a variable identifying the type of record
will be included. A new value of this identifier
will be used whenever there is a significant change
of the procedures used to generate that type of
record. Such changes include changes not only in
record layout, but also in instrument design, data
collection, and coding. The need for a new record
type is obvious when coding instructions or code
values are changed. However, even when such changes
consist only of the addition of new coding categories
to existing variables, a new record type should be
produced. Otherwise, analysts may not properly
interpret the absence of a particular response.

For example, several minor changes were made in
the coding categories and format of the incident
record of the National Crime Survey victimization
file. In order to interpret a "type I" record
properly, it is necessary to know what year the data
in the record represent. The record type and
collection period should be used to generate a set of
identifiers for incident records, identifying the

---

[4] A scaling factor allows fewer characters to be used
in representing a variable. For instance,
$50,000,000 can be represented as "50" if the
variable is defined as "Dollars in millions."

particular format used.

Record identification items. Each record in a
file submitted to the bureau must carry an
identification number unique to the data file. If no
existing variable will suffice as a unique
identifier, then a sequence number should be assigned
by the data supplier. Where a file contains more
than one type of record, each record shall carry a
variable identifying the type of record, as well as a
unique sequence number. Where the records in a file
represent a hierarchy or tree, a record will carry
unique identification sufficient to identify it and
its position in the hierarchy. In particular, it
should not be necessary to infer the location of a
record in a hierarchy solely from its position in the
file.

For example, consider a file consisting of
household, person, and incident records. Each record
in the file must begin with the same four data items:
a record type indicator, a household identifier, a
person identifier, and an incident identifier. An
incident record will carry a type and an incident
identifier, as well as the identifiers of the person
and household to which it belongs. A person record
will carry a type identifier, the identifier of the
household to which it belongs, its own identifier,
and a dummy incident identifier. Record identifiers
should be positive integers. Dummy identifiers
should be fields of zeros. Blank identification
variables are not permitted.

In general, identifiers of lower-level records
need be unique only within level, since concatenating
identification variables generates a unique
identifier.[5] In some files, there will be more than
one type of record at the same level of the
hierarchy. Questions as to whether sequence numbers
should be unique within record type or within levels
should be resolved by agreement between the data
supplier and the bureau before the file is generated.

---

[5] To concatenate identifiers is to string them
together into a single identifier. Consider a file
of households with people in them. Suppose that each
household has a three-digit identifier which is uni-
que within the file, and each person has a two-digit
identifier which is unique within his or her house-
hold. Then person number 4 in household number 207
is uniquely identified as person number 20704.

The rationale underlying the assignment of a
unique identifier to each record is that users of a
data file should be able to perform arbitrary sorts
and to subset the file without requiring the use of
any facilities other than sort program and a file-
copying utility.

Standardization of data codes

Where possible, standard data codes should be
used. There are no universal standards, but it is
often possible to find an existing and well-
constructed set of codes that is appropriate. There
are two reasons for emphasizing the adoption of
existing coding schemes. The first reason is that
the adoption of standard coding schemes facilitates
the linking of files. The second, and more
important, reason is that the development of coding
schemes for variables with large numbers of
categories is a considerable project in itself. It
is risky and uneconomical for a project to undertake
the development of its own coding scheme for such
things as geographical locations, occupations,
criminal offenses, mental diseases, etc. The choice
of which "standard" coding scheme is used is not so
important as that **ad hoc** coding schemes not be used
if equally good coding schemes are already in use.

Data suppliers subject to this standard are
encouraged to consult the following sources of infor-
mation on standard coding schemes.

Federal Information Processing Standards (FIPS).
The Federal Information Processing Standards
published by the National Bureau of Standards are the
official Federal standards for several coding
schemes. FIPS coding schemes should be considered
for geographic area variables such as states,
counties, Standard Metropolitan Statistical Areas
(SMSAs), populated places, etc. FIPS incorporate
several American National Standards Institutes (ANSI)
standards which have been adopted for official
Federal use. Data suppliers can obtain the current
list of FIPS publications, NBS Publications List 58,
from the Office of Standards Administration,
Institute for Computer Sciences and Technology,
National Bureau of Standards, Washington,
D. C. 20234, telephine (301) 921-3157.

Bureau of the Census. The Bureau of the Census
has its own set of geographical area codes as well as

a continually revised set of occupation and industrial codes. Data suppliers whose files are related to Census products should communicate with the Data User Services Division, Bureau of the Census, Washington, D. C. 20233 to receive copies of the latest Census coding schemes.

National Crime Information Center. The Federal Bureau of Investigation's National Crime Information Center (NCIC) maintains a list of Uniform Offense Codes and a list of codes for more than 60,000 criminal justice agencies. Data suppliers requiring copies of these coding schemes should communicate with the National Crime Information Center, Federal Bureau of Investigation, Washington, D. C. 20535.

Office of Federal Statistical Policy and Standards. The Office of Federal Statistical Policy and Standards, U. S. Department of Commerce, serves as a central clearinghouse for data representations and coding schemes. Data suppliers who are unable to locate a satisfactory data coding scheme are encouraged to communicate with the Office of Federal Statistical Policy and Standards, U. S. Department of Commerce, Washington, D. C. 20230, telephone (202) 673-7956.

Abstract

When data files are submitted to the archive, the producer must provide both the bureau and the archive with an abstract of the file's form and contents. The abstract must include bibliographic and technical information sufficient to meet the requirements of the Office of Federal Statistical Policy and Standards concerning public use data files released by Federal Agencies. The writing of the abstract is entirely the responsibility of the file's producer. The abstract must contain all elements listed below which are relevant to the form and content of the file.

1.   Date of the abstract.

2.   Unique file number (producer's number).

3.   **Bibliographic citation.**

     3.1.   **Title.** The title should be descriptive of the contents of the data file and should be followed by the date of the data program, if applicable. The title may be changed by the bureau if it does not clearly delineate the file from other files.

     3.2.   **Statement of responsibility.** The agency or agencies responsible for the file's substantive content. If a contractor produces the file, both the contractor and the contracting agency should be included.

     3.3.   **Place of production.** The formal mailing address of the organization which produced the file.

     3.4.   **Producer.** The organization which actually produced the file. This entry is usually redundant with the statement of responsibility.

     3.5.   **Date of production.** The date the file became operational in a computerized form and available for possible inclusion in the archive.

4.   **General description.**

     4.1.   **Type of file.** Types of files include text files, aggregate data files , microdata files, etc.

     4.2.   **Universe description.** A description of the population covered by the data in the file, including what, if any, sampling method was used.

     4.3.   **Subject matter description.** A description of the types of variables included on the file and any other comments about the purpose of the data collection, the scope of the study, or any special characteristics of the study.

5. **Time coverage.** The time period to which the data in the file refer, including dates of collection.

6. **Geographic coverage.** Information concerning the total geographic scope of the file and additional levels of geographic coding provided within the file.

7. **Technical description.** A brief description of the file structure (hierarchical, rectangular, etc.), the file size (the approximate number of logical records in the file), and any information about the format of the file (SPSS, SAS) or dependency on specific computers or software.

8. **Reference or related materials.** The bibliographic citation of all materials that relate to the machine-readable file or its technical aspects such as the technical documentation or that include analysis, results, or tables derived from the data.

9. **Related machine-readable files.** The bibliographic citation of other related data or software files that are available to the public or qualified users.

10. **Office of Management and Budget clearance number.** The number given by OMB to any form(s) or plan(s) used in the data collection. Please indicate if OMB clearance was not required.

11. **Terms of availability.** Any restrictions on access to the file such as privacy certification and a contact person(s) in the producing organization who can discuss the substantive and computing aspects of the file. The name, full address, and telephone number of the contact person(s) must be included.

Documentation

Each tape submitted for possible inclusion in the archive must be accompanied by documentation giving the physical characteristics of the volume and files as well as the logical composition of each type of record. Wherever possible, the documentation should be in machine-readable form and supplied both

in hard copy and as a file on the tape. Machine-readable documentation is preferred because it prevents the separation of documentation from data and because the physical quality of the documentation will not be degraded by repeated copying.

Machine-readable documentation should be in printer image. If a document processing program is used to format the source text, it is requested that the source text of the documentation and a reference to the text processing software be included as well. (This standard makes no specification of any particular document processing program, nor does it require that such a program be used at all.) The minimal documentation of a data file consists of a tape volume table of contents, a character and octal or hexadecimal dump of a sample of records of each file, information sufficient to construct a bibliographic citation to the machine-readable data file, an abstract of the file's form and content, and a minimal codebook.

Where appropriate, data files should be accompanied by copies of the original collection instruments, including survey questionnaires and interview schedules. Copies of editing and coding instructions used in the creating the data file should also be included.

Tape Table Of Contents

The tape volume table of contents listing should include all information from the volume label and from the file labels. Information should be in an easy-to-read form, rather than a dump of the text of the labels. The following figure is an example of a table volume table of contents listing.

If possible, the tape table of contents should be produced by a program which also verifies the readability of the tape. If no tape listing program is available, then the tape table of contents should be produced manually, using information from the job which produced the tape or a dump of the tape. (Since tape listing programs are becoming increasingly available, it is suggested that data centers without such programs attempt to acquire them.)

Minimal Codebook

The codebook included with the file must contain at least the following information for each variable:

1. A reference number.

2. An unambiguous name for the item.

3. A textual description of the item, or the text of the question, if from a questionnaire.

4. The starting location, width, location of implicit decimal point, or scale factor.

5. Missing data codes and their meanings.

6. The mode in which the variable is represented, i.e., numeric character, alphanumeric string, floating-point binary, etc.

The codebook must also contain a list of the valid values for categorical items and valid ranges for continuous items. Missing data codes must be documented in the same fashion as other values and not left implicit.

Frequency Tables

A frequency distribution for each categorical variable must accompany each file submitted for possible inclusion in the archive. The mean, standard deviation, range and number of cases of continuous variables should also accompany the file. Values that fall outside of those defined in the codebook should be annotated if they cannot be corrected.

```
TAPE NAME = *RSW009*    24 MAY   1977    12:20:06


IBM-LABELED 6250-BPI    9TP VOLUME=RSW009 OWNER=CAC,U-ILL   RACK#=C4524
LP=ON  BLK=ON  RING=OUT DTCHK=ON   RETRY=10

FILE                     BLOCK RECORD TAPELTH  RECORD             BLOCK LTH    CREATED   EXPIRES      USER     BATCH
  # DATA SET NAME        COUNT  COUNT (FEET)   FORMAT             AV.   MAX.  DD MMM YY DD MMM YY  I.D.      RECEIPT#

  1 NCS73.NAT.CQ1        1233 125858  283.07  VB(15250,305)      15141 15250 23 MAR'77             SGDA
  2 NCS73.NAT.CQ2        1234 126442  283.17  VB(15250,305)      15135 15250 23 MAR'77             SGDA
  3 NCS73.NAT.CQ3        1187 123279  272.60  VB(15250,305)      15146 15250 23 MAR'77             SGDA      691127
  4 NCS73.NAT.CQ4        1212 125436  278.21  VB(15250,305)      15139 15250 23 MAR'77             SGDA      691127
  5 NCS74.NAT.CQ1        1086 112135  249.39  VB(15250,305)      15139 15250 23 MAR'77             SGDA      691127
  6 NCS74.NAT.CQ2        1083 112086  248.84  VB(15250,305)      15148 15250 23 MAR'77             SGDA      691127
  7 NCS74.NAT.CQ3        1072 111090  246.18  VB(15250,305)      15138 15250 23 MAR'77             SGDA      691127
  8 NCS74.NAT.CQ4        1100 113455  252.62  VB(15250,305)      15141 15250 23 MAR'77             SGDA      691127

TOTAL TAPE LENGTH = 2114.07 FEET

<*><*><*>   END OF TAPE   <*><*><*>
```

Figure 1: A tape volume table of contents.

NCJRS REGISTRATION

The National Criminal Justice Reference Service (NCJRS) abstracts documents published in the criminal justice field. Persons who are registered with the Reference Service receive announcements of documents in their stated fields of interest and order forms for free copies of Bureau of Justice Statistics publications. If you are not registered with the Reference Service, and wish to be, please provide your name and mailing address below and check the appropriate box.

Name		Telephone	
		( )	☐ Please send me a NCJRS registration form.
Number and street			
			☐ Please send me the reports listed below.
City	State	ZIP Code	

(Fold here)

U.S. DEPARTMENT OF JUSTICE
Bureau of Justice Statistics
Washington, D.C. 20531

```
PLACE
STAMP
HERE
```

User Services Department 2
National Criminal Justice Reference Service
Bureau of Justice Statistics
U.S. Department of Justice
Box 6000
Rockville, Maryland 20850

(Fold here)

If you wish to receive copies of any of the Bureau of Justice Statistics Reports listed on the reverse side, please list them below.

_____

_____

_____

_____

Bureau of Justice Statistics Reports

Single copies are available at no charge from the National Criminal Justice Reference Service, Box 6000, Rockville, Md. 20850. Multiple copies are for sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

National Crime Survey:
Criminal Victimization in the United States (annual):
  Summary Findings of 1978-79 Changes in Crime and of Trends Since 1973, NCJ-62993
  A Description of Trends from 1973 to 1978, NCJ-66716
  1978 (final report), NCJ-66480
  1977, NCJ-58725
  1976, NCJ-49543
  1975, NCJ-44593
  1974, NCJ-39467
  1973, NCJ-34732
The Cost of Negligence: Losses from Preventable Household Burglaries, NCJ-53527
The Hispanic Victim: Advance Report, NCJ-67706
Intimate Victims: A Study of Violence Among Friends and Relatives, NCJ-62319
Crime and Seasonality, NCJ-64818
Criminal Victimization of New York State Residents, 1974-77, NCJ-66481
Criminal Victimization of California Residents, 1974-77, NCJ-70944
Indicators of Crime and Criminal Justice: Quantitative Studies, NCJ-62349
Criminal Victimization Surveys in 13 American cities (summary report, 1 vol.), NCJ-18471
  Boston, NCJ-34818
  Buffalo, NCJ-34820
  Cincinnati, NCJ-34819
  Houston, NCJ-34821
  Miami, NCJ-34822
  Milwaukee, NCJ-34823
  Minneapolis, NCJ-34824
  New Orleans, NCJ-34825
  Oakland, NCJ-34826
  Pittsburgh, NCJ-34827
  San Diego, NCJ-34828
  San Francisco, NCJ-34829
  Washington, D.C., NCJ-34830
Public Attitudes About Crime (13 vols.):
  Boston, NCJ-46235
  Buffalo, NCJ-46236
  Cincinnati, NCJ-46237
  Houston, NCJ-46238
  Miami, NCJ-46239
  Milwaukee, NCJ-46240
  Minneapolis, NCJ-46241
  New Orleans, NCJ-46242
  Oakland, NCJ-46243
  Pittsburgh, NCJ-46244
  San Diego, NCJ-46245
  San Francisco, NCJ-46246
  Washington, D.C., NCJ-46247
Criminal Victimization Surveys in Chicago, Detroit, Los Angeles, New York, and Philadelphia: A Comparison of 1972 and 1974 Findings, NCJ-36360
Criminal Victimization Surveys in Eight American Cities: A Comparison of 1971/72 and 1974/75 Findings—National Crime Surveys in Atlanta, Baltimore, Cleveland, Dallas, Denver, Newark, Portland, and St. Louis, NCJ-36361

Applications of the National Crime Survey Victimization and Attitude Data:
Public Opinion About Crime: The Attitudes of Victims and Nonvictims in Selected Cities, NCJ-41336
The Police and Public Opinion: An Analysis of Victimization and Attitude Data from 13 American Cities, NCJ-42018
Compensating Victims of Violent Crime: Potential Costs and Coverage of a National Program, NCJ-43387

Local Victim Surveys: A Review of the Issues, NCJ-39973
An Introduction to the National Crime Survey, NCJ-43732
Crime Against Persons in Urban, Suburban, and Rural Areas: A Comparative Analysis of Victimization Rates, NCJ-53551
Rape Victimization in 26 American Cities, NCJ-55878
Criminal Victimization in Urban Schools, NCJ-56396
Restitution to Victims of Personal and Household Crimes, NCJ-72770

National Prisoner Statistics:
Capital Punishment (annual):
  1979, NCJ-70945
Prisoners in State and Federal Institutions on December 31:
  1979, NCJ-73719
Census of State Correctional Facilities, 1974 advance report, NCJ-25642
Profile of State Prison Inmates: Sociodemographic Findings from the 1974 Survey of Inmates of State Correctional Facilities, NCJ-58257
Census of Prisoners in State Correctional Facilities, 1973, NCJ-34729
Census of Jails and Survey of Jail Inmates, 1978, preliminary report, NCJ-55172
Profile of Inmates of Local Jails: Sociodemographic Findings from the 1978 Survey of Inmates of Local Jails, NCJ-65412
The Nation's Jails: A report on the census of jails from the 1972 Survey of Inmates of Local Jails, NCJ-19067

Uniform Parole Reports:
Parole in the United States (annual):
  1979, NCJ-69562
  1978, NCJ-58722
  1976 and 1977, NCJ-49702
A National Survey of Parole-Related Legislation Enacted During the 1979 Legislative Session, NCJ-64218
Characteristics of the Parole Population, 1978, NCJ-66479

Children in Custody: Juvenile Detention and Correctional Facility Census
1977 advance report:
  Census of Public Juvenile Facilities, NCJ-60967
  Census of Private Juvenile Facilities, NCJ-60968
1975 (final report), NCJ-58139
1974, NCJ-57946
1973, NCJ-44777
1971, NCJ-13403

Myths and Realities About Crime: A Nontechnical Presentation of Selected Information from the National Prisoner Statistics Program and the National Crime Survey, NCJ-46249

State and Local Probation and Parole Systems, NCJ-41335
State and Local Prosecution and Civil Attorney Systems, NCJ-41334
National Survey of Court Organization:
  1977 Supplement to State Judicial Systems, NCJ-40022
  1975 Supplement to State Judicial Systems, NCJ-29433
  1971 (full report), NCJ-11427
State Court Model Statistical Dictionary, NCJ-62320
State Court Caseload Statistics:
  The State of the Art, NCJ-46934
  Annual Report, 1975, NCJ-51885
  Annual Report, 1976, NCJ-56599
A Cross-City Comparison of Felony Case Processing, NCJ-55171

Trends in Expenditure and Employment Data for the Criminal Justice System, 1971-77 (annual), NCJ-57463
Expenditure and Employment Data for the Criminal Justice System (annual)
  1978 Summary Report, NCJ-66483
  1978 final report, NCJ-66482
  1977 final report, NCJ-53206

Dictionary of Criminal Justice Data Terminology: Terms and Definitions Proposed for Interstate and National Data Collection and Exchange, NCJ-36747

Justice Agencies in the U.S.: Summary Report of the National Justice Agency List, NCJ-65560
Criminal Justice Agencies in Region
  1: Conn., Maine, Mass., N.H., R.I., Vt., NCJ-17930
  2: N.J., N.Y., NCJ-17931
  3: Del., D.C., Md., Pa., Va., W.Va., NCJ-17932
  4: Ala., Ga., Fla., Ky., Miss., N.C., S.C., Tenn., NCJ-17933
  5: Ill., Ind., Mich., Minn., Ohio, Wis., NCJ-17934
  6: Ark., La., N.Mex., Okla., Tex., NCJ-17935
  7: Iowa, Kans., Mo., Nebr., NCJ-17936
  8: Colo., Mont., N.Dak., S.Dak., Utah, Wyo., NCJ-17937
  9: Ariz., Calif., Hawaii, Nev., NCJ-15151
  10: Alaska, Idaho, Oreg., Wash., NCJ-17938

Utilization of Criminal Justice Statistics Project:
Sourcebook of Criminal Justice Statistics 1980 (annual), NCJ-71096
Public Opinion Regarding Crime, Criminal Justice, and Related Topics, NCJ-17419
New Directions in Processing of Juvenile Offenders: The Denver Model, NCJ-17420
Who Gets Detained? An Empirical Analysis of the Pre-Adjudicatory Detention of Juveniles in Denver, NCJ-17417
Juvenile Dispositions: Social and Legal Factors Related to the Processing of Denver Delinquency Cases, NCJ-17418
Offender-Based Transaction Statistics: New Directions in Data Collection and Reporting, NCJ-29645
Sentencing of California Felony Offenders, NCJ-29646
The Judicial Processing of Assault and Burglary Offenders in Selected California Counties, NCJ-29644
Pre-Adjudicatory Detention in Three Juvenile Courts, NCJ-34730
Delinquency Dispositions: An Empirical Analysis of Processing Decisions in Three Juvenile Courts, NCJ-34734
The Patterns and Distribution of Assault Incident Characteristics Among Social Areas, NCJ-40025
Patterns of Robbery Characteristics and Their Occurrence Among Social Areas, NCJ-40026
Crime-Specific Analysis:
  The Characteristics of Burglary Incidents, NCJ-42093
  An Empirical Examination of Burglary Offender Characteristics, NCJ-43131
  An Empirical Examination of Burglary Offenders and Offense Characteristics, NCJ-42476
Sources of National Criminal Justice Statistics: An Annotated Bibliography, NCJ-45006
Federal Criminal Sentencing: Perspectives of Analysis and a Design for Research, NCJ-33683
Variations in Federal Criminal Sentences: A Statistical Assessment at the National Level, NCJ-33684
Federal Sentencing Patterns: A Study of Geographical Variations, NCJ-33685
Predicting Sentences in Federal Courts: The Feasibility of a National Sentencing Policy, NCJ-33686

Technical Standards for Machine-Readable Data