Center for Cybercrime Studies

John Jay College of Criminal Justice
Mathematics & Computer Science Department
445 West 59th Street
New York, NY 10019

Technical Report

Exploring NIBRS with a Relational Database

Eman Abdu, Douglas Salane & Peter Shenkin

August 2011

Exploring NIBRS Data with a Relational Database

Eman Abdu, Douglas E. Salane and Peter Shenkin

Center for Cybercrime Studies
Department of Mathematics & Computer Science
John Jay College of Criminal Justice

August 8, 2011

Abstract

The FBI's National incident-based crime report system (NIBRS) data collection provides exceptional opportunities to understand the situational and demographic factors associated with crime. This paper describes a relational database implementation of NIBRS. The paper provides an introduction to NIBRS data structure, discusses current tools available for analyzing NIBRS data, and explores the benefits of analyzing NIBRS with a relational database. The paper examines how NIBRS data is changing over time and some of the data quality issues that arise from missing or infrequently provided data. The paper also explores the use of the relational database and spreadsheet Pivot Tables to make NIBRS analysis capabilities available to analysts for specific studies.

1. Introduction

In the 1980s, the FBI, as part of its Uniform Crime Reporting (UCR) program, developed the National Incident-Based Reporting System (NIBRS) (Federal Bureau of Investigation, 2009). The NIBRS collects detailed information on crime incidents reported to law enforcement agencies and makes that information available for crime research and analysis. Unlike the FBI's UCR Summary Reporting System, which provides mainly aggregate crime statistics for eight serious forms of crime[1], the NIBRS collects detailed information on 46 specific crimes in 22 offense categories[2]. For a given crime incident, NIBRS records up to 57 data elements that provide information on the offense, victim, offender, property involved, arrestee and the incident itself, e.g., the date and time of the incident as well as the reporting agency. With NIBRS, researchers typically have information on the situational context of the incident, victim-offender demographics, victim-offender relationships, and incident clearance. A given NIBRS crime incident record can include information on multiple offenses, up to ten victims and offenders, and multiple arrestees. In addition, NIBRS also provides information on the reporting agency and even population information for the reporting agency's locale. Currently NIBRS data for the years 1991 through 2008 are available for download from the National Archive of Criminal Justice Data (NAJCD, 2010).

[1] The UCR program provides counts for murder and nonnegligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny-theft, motor vehicle theft, and arson.

[2] The 46 offenses for which NIBRS *provides* detailed information are known in NIBRS as Group A incidents. NIBRS also collects information on a set of less serious non-violent, victimless offenses known as Group B offenses for which mainly arrestee data is recorded.

NIBRS allows a crime researcher to analyze the social, economic and demographic factors associated with crime. The paper by (Maxfield, 1999) describes the opportunities NIBRS presents for research as well as the potential pitfalls. With NIBRS, criminal justice researchers have examined aspects of crime such as the relationships between victims and offenders, weapons used in various types of crimes, crime trends over time and in different localities, and the prevalence of certain types of crime. NIBRS is one of the few crime databases that provide the information needed for the study of crime clearance and its relation to race, age and other demographic factors. The report (Roberts, 2009) offers a comprehensive survey of the use of NIBRS for study of crime clearance, disaggregation of crime rates by situational context, and outcomes of violent encounters.

Despite its tremendous potential, researchers have noted several difficulties with NIBRS data. These include the frequency of missing incident data and the lack of agency participation, especially in larger municipalities, which could lead to biased analyses (Addington, 2008). Other concerns include lack of incident-level variables that provide educational, occupational and precise location information, i.e., the location of the incident within the reporting agency, making it difficult to determine a socio-economic context. Nonetheless, the use of NIBRS for reporting crime data is increasing and the quality appears to be improving as agencies continue to adopt NIBRS based operational record management systems that provide incident-based crime data to the UCR program through NIBRS. (See Appendix I for a discussion.) In an era when crime rates are holding steady, the number of crime incidents in the NIBRS collection has grown 51% in just the last three years for which NIBRS data is available (2006 – 2008). The high rate of increase is due primarily to increased agency reporting.

Key impediments to research with NIBRS data have been the complexity and size of the data set as well as limited tools for data extraction and analysis. Despite dramatic improvements during the past 10 years in the computer processing and storage capabilities available to researchers, the size of the data set still poses problems. Time series studies are now possible with NIBRS but the entirel data set can require significant storage and processing capabilities. Moreover, NIBRS data are available for a particular year and multiyear analysis requires processing multiple files. In addition, the 57 different data elements associated with an incident are grouped into 6 different categories or segments: Administrative, Offense, Property, Victim, Offender, and Arrestee. Examining, for example, victim-offender relationships requires linking the data in the victim and offender segments together. Tools available at the National Archive of Criminal Justice Data, described briefly in Section 3, mitigate some of the complexity of dealing with NIBRS data. These tools give criminal justice researchers significantly increased access to NIBRS data and provide some statistical analysis capabilities. However, they provide limited capabilities for exploring relationships in NIBRS data. They are not intended for multiyear analysis. In addition, database analysts are not able to use them to provide schema and customized views of the data tailored for particular research investigations, which researchers could then further analyze by performing iterative queries. Moreover, modern relational databases provide a wide range of built-in facilities that take advantage of the structure of the data for efficient processing and analysis.

In this paper we describe a relational database implementation of NIBRS data that mirrors the structure of NIBRS. The relational implementation allows an analyst to explore NIBRS data in ways not possible with current facilities. Our implementation, which contains data from the years 1995 to 2008, allows us to examine various NIBRS data quality issues including the frequency of missing data as well as the extent of multiple records and their affect on analysis. Furthermore, the database allows us to prepare views of the data and schema that facilitate analysis for non data processing experts. For example, we illustrate simple joins that create a view, or an intermediate table, that has the required information for victim-offender analyses. We also show how a customized view can be imported into a

spreadsheet where modern pivot tables provide capabilities similar to data cubes. In addition, to multi segment analyses, our database implementation facilitates multiyear analysis.

The following is an outline of the paper. In section 2, we give a brief overview of the structure of NIBRS data. In section 3, we review current sources of NIBRS data and analysis tools available to researchers. We describe in section 4 our relational database implementation and tools we use for updating our implementation when new NIBRS data becomes available. In section 5, we use the relational database to examine various aspects of NIBRS data and present several sample studies. In section 6, we use the relational database to extract data needed to study the use of a computer in various offenses. We discuss how the extracted data can be imported into a spreadsheet to provide analysts with capabilities similar to those of a so-called Data Cube. Here we use the capabilities to view the offender counts disaggregated by race, gender and age. We conclude with a brief description of future work.

2. A brief review of NIBRS structure

The organization of NIBRS data and the potential complexities of analyzing it have been covered in many papers and reports. See for example (Akiyama & Nolan, 1999). One of the best sources of detailed information on NIBRS data structure is the most recent code book that is available at National Archive of Criminal Justice Data (NACJD) (Federal Bureau of Investigation, 2008). A good overview of NIBRS structure appears on the web site (NACJD, 2010). In this section we simply give an overview of the basic NIBRS structure that is essential for understanding and using our relational database implementation. The section also provides an overview of the types of detailed information in NIBRS and gives an indication of some of the data processing questions that NIBRS presents.

As noted, the NIBRS classifies crime incidents into two different categories called Group A and Group B offenses. Group A NIBRS incidents include crimes mainly against person and property and result from an incident being reported to a law enforcement agency. Group B offenses consist of crimes against society such as drunkenness, driving under the influence, and disorderly conduct. Group B crimes typically result from an arrest. There are 46 Group A offenses and 11 Group B offenses.

NIBRS handles Group A and B incidents quite differently. Group A incidents include information on one or more offenders, victims, and offenses associated with the incident. A Group A incident may also have information on one or more arrestees and, depending on the offense(s), the property involved in the incident. Group B incidents mainly provide demographic information for the arrestee as well as information on the offense, possible weapons involved, and the arrest itself. NIBRS data from 1995 to 2008 consists of approximately 44.1 million Group A incidents and about 14.6 million Group B incidents. This report focuses primarily on Group A incidents since it is for these incidents that a relational database implementation provides the greatest utility.

For Group A incidents, NIBRS data organization centers around a crime incident reported to a law enforcement agency. Each NIBRS Group A incident consists of 57 data elements that provide specific information about the incident. NIBRS groups the data elements into six different categories or so-called segments: Administrative, Offense, Property, Victim, Offender, and Arrestee segments. A single record in the Administrative Segment provides information on a given incident, for example, the date and hour, the reporting agency, and clearance information. In NIBRS, information on the offenses, property involved, victims, offenders and arrestees appears in the five other segments. For example, information on one offense is recorded in a single record in the Offense Segment. The incident record in

the Administrative Segment provides counts of how many offense, property, offender, victim, and arrestee records are associated with the incident. Each segment record includes the Originating Agency Identifier (ORI) and incident number, which together identify all segments records associated with a particular incident. In this respect NIBRS is very much like a relational database. Appendix II lists the six main Group A incident segments and associated fields as they appear in our implementation of the NIBRS database.

Group B offenses appear in a separate segment called the Group B Arrest Segment. Each record corresponds to an arrest for a Group B offense. Group B information includes primarily characteristics of the arrestee: age, sex, race, ethnicity and resident status. Group B offense records also contain an ORI and incident number. There may be multiple Group B arrest records with the same ORI and incident number if more than one person is arrested for the offense.

In addition to the seven segments just described, NIBRS provides six additional segments. There are three so-called Batch Header Segments that provide information on the originating agency. The ORI number, which identifies the agency, is the identifying field in each of these three segments. These segments contain information on the agency's location, city and county the agency serves, population indicators, and even NIBRS submission history. Three more segments, the Windows Exceptionally Cleared, Windows Recovered Property, and Window Arrestee Segments, provide a mechanism for agencies to update incident information if new additional information, e.g., clearance, arrest or property recovered information, becomes available after an initial incident report is submitted. Indeed, a consideration in NIBRS design was the ability to update incident information after an initial submission. In section 4, we discuss the use of duplicate records in subsequent years, which also is an important mechanism for updating NIBRS information.

One way NIBRS differs from a traditional normalized database is the use of multiple entries for a data element. For example, one data element for an incident is the offense code. A NIBRS record in the victim segment includes up to 10 different offenses against a victim. A victim segment record provides space for these 10 different offense codes even if most are blank. In addition, a victim segment record lists the victim-offender relationships for up to 10 different offenders. The victim segment record also provides space for up to 5 different victim injuries. Offense segment records can contain up to 3 codes that categorize the type of criminal activity associated with the offense and 3 codes that indicate the type of force/weapon involved. An offense record also has 3 fields that indicate whether the offense involved a computer, drugs, alcohol, or some combination of these. Allowing for multiple responses for data elements actually facilitates NIBRS data processing. In Appendix II, the listings of the fields in the main NIBRS segment s indicate the multi response data elements.

One of our interests has been the frequency with which multi response data is missing in NIBRS and how that frequency varies over time. We suspect that as agencies increasingly deploy NIBRS based operational systems, data collection of these types of attributes will increase. In section 5, we present some preliminary findings on the frequency and quality of this type of NIBRS data.

Analysis of Group A incidents presents several challenges. Each incident may involve one or more offenses, victims, arrestees, properties, etc. Depending on the study, for example, if the researcher is counting incidents, one or more of these segments must be linked to the associated administrative record. If a count of victims is the unit of analysis, one or more segments must be linked to the victim record. As noted, further complicating NIBRS processing, certain data elements allow for the recording of multiple responses, for example, different weapons used, different victim injuries, the

relationship of a victim to different offenders. The complexity of NIBRS data has frequently led researchers to restrict their attention to single segment analysis or to limit analysis to one additional segment. Frequently, researchers ignore multiple offenders or victims associated with an incident. Some of the difficulties of dealing with NIBRS complexity are summarized in (Chilton & Regoeczi, 2005) and (Akiyam, October 1997). Researchers need additional tools and effective methodologies for extracting and analyzing complex incident based data.

3. Sources of NIBRS data and some current tools

Researchers involved in NIBRS realized early on that the size and complexity of NIBRS database would require data processing skills that typically exceeded those of most data analysts. The National Archive of Criminal Justice Data (NACJD) at the Inter-university Consortium for Political and Social Research (ICPSR) has played the lead role in packaging NIBRS data and making it available to the Criminal Justice research community (Dunn & Zelenock, 1999). Indeed, our own work relies on NIBRS data made available for download from ICPSR. Here we briefly describe the efforts of ICPSR and others to make NIBRS data available and provide analysis capabilities.

The FBI releases NIBRS data for a given year as one large rectangular file in a format similar to that of a spreadsheet. As discussed, NIBRS data elements are organized by segments. In the FBI released file, each row corresponds to a record in a segment and the columns correspond to data elements. The records in a segment are stored in contiguous rows. Most data elements (a cell in the spreadsheet) contain a code, typically, one or two characters in length, which is assigned by the FBI and described in a codebook released with the data. The FBI also assigns an agency identifier and an incident number, which appear in each segment record and allow segments associated with a particular incident to be linked together. Data processing is complicated by the size of the segments and the need to link segments. For example, for 2008 the Administrative Segment occupies over 5 million rows while the Victim Segment, the largest segment, occupies 5.5 million rows.

Most researchers obtain NIBRS data from the ICPSR's NACJD web site[3]. After checking FBI released data for consistency[4], ICPSR separates the data in the FBI NIBRS file into 13 separate files and makes these available for download. ICPSR also adds a variable to each segment record that indicates the number of segment records that are associated with a particular incident. For example, if an incident has multiple victim records, the first victim record, identified by a victim sequence number, would contain a count of the number of victim segments associated with the incident. In addition, ICPSR provides a modified codebook that details the structure of the files as prepared by ICPSR. Single segment analysis is facilitated by the fact the NAJCD provides SPSS and SAS scripts to downloaded NIBRS data into those applications. In the early days of NIBRS, single segment analysis was the norm.

In order to facilitate multi segment analysis, NAJCD also provides a set of NIBRS data files called Extract Files that contain merged segments (NACJD, 2008). The Extract Files are based on the desired unit of analysis such as incident, victim or offense. For example, the analyst selects the Administrative-level Extract file if a count of incidents is desired. A record in the Administrative-level Extract file includes a given incident merged with the records from the five other segments associated with the given incident. Similarly, a Victim-level Extract file would include Victim records merged with the five

[3] As of this writing NIBRS data for the years 1991 to 2008 were available.

[4] For example, ICPSR checks that codes in that appear in the data are consistent with the descriptions given in the NIBRS codebook.

other associated segments. The extract file has the same number of records as the unit of count upon which it is based. In order to keep extract files from becoming unwieldy, the ICPSR only merges up to three segment records. As the ICPSR points out and our analysis in the next section confirms, 99% of all NIBRS incidents have 3 or fewer such records.

The National Center for Juvenile Justice (NCJJ, 2010) has used the 2008 Victim-Level Extract files to produce a modified extract file suitable for studying victims of domestic violence. First, it culled the original ICPSR extract file to include only victims of domestic violence. NCJJ also defined a person offense[5] hierarchy and only included those records where the most serious offense was a person offense. In addition, NCJJ eliminated states that reported fewer than 1000 domestic violence victims. This reduced a data set of about 5.5 million records to about 550 thousand records. If the data were in a relational database, such a customized view of the Extract file would be easy to produce with an SQL query. The NCJJ site also provides on-line analysis capabilities for this domestic violence data set.

There have been several other attempts to make on-line analysis capabilities available for NIBRS data. For over ten years the NACJD web site has used the Survey Documentation Analysis system (SDA, 2011) to allow users to analyze NIBRS data within a single segment and single year. Statistical analysis capabilities include frequency or cross tabulations, comparison of means, correlation matrix, and multiple regression. The addition of drop down menus with code definitions eliminates the need to specify codes and simplifies use of the system. Recently, the NACJD site has made the SDA System available for use on the extract files, which make possible multi segment analysis. The SDA interface, however, limits the ability to extract relations in the data and explore NIBRS for missing data. Moreover, it was not designed to provide customized views and schema often requested by crime researchers.

A second site that provides NIBRS analysis capabilities is The National Consortium on Violence Research (NCOVR). NCOVR provides web access to a hierarchical multidimensional data cube generated from a relational database implementation of NIBRS. No query capabilities are available. The dimensions reflect the organization of NIBRS data into segments. The cubes allow users to select and filter NIBRS variables and count incidents. The NCOVR site allows users to aggregate data from regions of the country down to the local reporting agency. The NCOVR site appears to have stopped supporting NIBRS as data are only available for the years 1998 through 2004. The report by (Edgar, Cohen, & Porter, 2004) provides an overview of the crime data resources available at NCOVR including the NIBRS.

In addition, to the aforementioned sites, the FBI is also plans to make simple on-line query analysis available for NIBRS data in 2013.

Both NCOVR and NACJD provide valuable services, but offer limited analysis capabilities for NIBRS data. NCOVR provides essentially multidimensional OLAP (MOLAP) processing in the form of data cubes. The cubes are restricted to NIBRS incident counts. Moreover, NCOVR cube dimensions focus on NIBRS segments relevant to violent incidents. All NCOVR access to NIBRS is through the data cubes; no query capabilities or relational OLAP (ROLAP) capabilities are available. NACJD extract files facilitate access to data in multiple segments. However, multiple response data elements make the extract files complex and significant processing is often required to extract desired information. Neither system makes it easy to explore relations in the data or perform iterative queries. Finally, both sites make

[5] Person offenses in this study include murder/nonnegligent manslaughter; kidnapping/abduction; forcible rape; forcible sodomy; sexual assault with an object; forcible fondling; robbery; aggravated assault; simple assault; and intimidation.

available a wide range of criminal justice data repositories and service a worldwide audience. Scalability severely limits the level of relational processing capabilities such sites can provide.

4. A relational database implementation

In this section we describe a relational database implementation of NIBRS data. Analysis facilities based on the SDA and data cubes allow end users to access and analyze NIBRS data. A relational database implementation, on the other hand, gives data processing specialists or anyone familiar with SQL, the opportunity to explore NIBRS data in ways not possible with systems designed primarily for end user analysis. For example, with a relational database it is easy to examine NIBRS for frequency of missing data, the extent to which multiple segments occur, and to extract relations in the data, for example, age versus multiple offenders and victims. (See section 5.) In addition, with a database implementation, specialists can create specialized views of the data for crime researchers, which then can be queried iteratively or exported into spreadsheets. Non-specialists can then employ features such as pivot tables in spreadsheets to analyze the data much as one would with a data cube. Finally, modern databases include a number of features that optimize processing of the data and help us verify the referential integrity of the data when adding newly released NIBRS data to our collection.

Our implementation mirrors the structure of the 13 NIBRS segment files provided by ICPSR (NACJD, 2010). ICPSR provides NIBRS data in 13 linked ASCII segment files, each of which is easily converted to a relational database table. Besides cleaning the data and checking it for referential integrity, ICPSR also includes in Victim, Offender, and Arrestee Segment records an attribute that indicates the number of such records associated with a given incident. Our implementation includes all NIBRS data available at ICPSR for the years 1995 through 2008 and thus facilitates multiyear analysis. A list of the attributes in each of the six main Group A segment tables in our implementation appears in Appendix II. Note that we have added the NIBRS release year as an attribute to all tables. This is required because incident records can be duplicated in subsequent years as discussed in the next section.

In our implementation, the Administrative Segment functions as a master table where each incident record is uniquely identified by the incident number, ORI code, NIBRS release year and incident date. Since all our tables include NIBRS data for multiple years, these four attributes together uniquely identify an incident in the Administrative Table and allow us to link information in other tables to a particular incident. Attributes in the Administrative Table, which appear in the original NIBRS data, indicate the total number of Victim, Offense, Offender, Property and Arrestee segments associated with a given incident. It is important to note that information for a NIBRS incident may be duplicated in a subsequent released year; i.e., information for one NIBRS incident may appear in two successive years. The only attribute that distinguishes the records is the NIBRS release year. The four aforementioned keys allow us to detect when incident information is duplicated in the data set. Researchers seem to be unaware of duplication in NIBRS; however, this occurs in a very small percentage of incidents, less than 1.5 % per year in the data we have analyzed.

We have implemented all other segments provided by ICPSR as tables and have added the required keys to link the information to an incident. In addition to the four keys mentioned, the victim, offender, and arrestee tables each contain a sequence number, which is needed to identify a record in those tables since each incident can include multiple records from these segments. For a given incident, for example, the sequence number in the victim table indicates the particular victim for which the victim

record provides information. Sequence numbers in the other tables serve the same purpose. Additional keys in the Offense and Property Segments distinguish multiple records associated with a single incident in those files. As noted, NIBRS uses a system of codes to indicate just about all data element values. For example, for a given incident, the offense table uses one of 58 different codes to indicate the particular offense value.

Besides the six main tables (Appendix II) that provide information on Group A incidents, NIBRS includes seven other segments which we include in our database implementation. NIBRS provides information on each ORI such as geographic, demographic, NIBRS submission and other organizational information. Three separate batch tables (Batch1, Batch2 and Batch 3) provide this data and in each the ORI number and NIBRS release year provide unique record identifiers. We include Batch files for all years for researchers who might be interested in studying, for example, how agency participation varies over time. In addition, Group B offenses are listed in a separate table that is not linked to other tables in the database. Here the ORI, Incident number, Incident date, NIBRS release date and arrestee sequence number are the attributes needed to distinguish records. The table includes an arrestee sequence number to distinguish different arrestees recorded for the same incident. Our implementation also includes as tables the three Windows Segments released with NIBRS. These segments provide information on exceptional clearance, recovered property and arrests that might be made after NIBRS data is submitted. Record counts of the three Windows tables, which appear in the next section, indicate this type of data is provided infrequently.

In addition to the 13 NIBRS tables built from NIBRS segments, our implementation includes 46 additional supporting tables. Since NIBRS data consist mainly of codes, the code descriptions found in the NIBRS codebook must be associated with these codes to understand the data. Our implementation provides tables that map codes to descriptions; supporting tables include UCR Offenses, Bias Motivations, Property Descriptions, Types of Activity, Relationships of Offenders and Victims, Type of Criminal Activity, Type of Weapon or Force Involved, etc. These additional tables are an important feature of our implementation as they are needed to enforce referential integrity, provide analysts with a quick description of the codes, and facilitate report generation. These tables can be viewed as our application's data dictionary and are derived from the NIBRS codebook.

In our implementation, we maintain~~ed~~ the NIBRS structure of multiple responses for certain data elements. Such implementations, although not characteristic of a normalized database, facilitate processing and limit the complexity of the database, especially for a data warehouse environment. As noted, the Victim, Arrestee, Offense and Property Segments (Tables) contain multiple response data elements. Multiple response data elements in the six main NIBRS segments are indicated in Appendix II.

A modern database[6] provides many optimizations that can take advantage of the structure of NIBRS data. For example, the indexing capabilities and caching schemes used can dramatically speed up processing. Even database joins, a computationally intensive operation, can take advantage of the fact that most incidents involve very few multiple segment records. Modern features such as partitioning can speed up a range of operations. In NIBRS data processing, for example, partitioning by year is very natural and allows those interested in specific years to avoid the large costs or processing tables that contain data for all years. For most studies, we prepare for researchers pre-computed data views (or tables) that support subsequent queries and require no further joins. Finally, modern database systems

[6] Our current implementation uses Oracle 11g Enterprise Edition, Release 11.2.0.1.0, 64 Bit Production

provide a host of add on tools for data mining, data warehousing and building custom applications, which are optimized to work with the database.

Another significant effort was the building of facilities to update the database when additional NIBRS data is made available by NACJD. We have developed scripts that add NAJCD segment files to our existing database. The database employs the supporting tables to enforce the referential integrity of the database when data is added. For example, if the FBI adds codes to NIBRS, the supporting tables must be updated. Indeed the referential integrity check that occurs when the data base is updated is a computationally intensive operation.(HP Proliant ML 370G5 Server with 8 Giga Bytes of RAM).

5. Understanding NIBRS data: A glimpse of what's there

In this section we use the relational database implementation to examine NIBRS data. We examine the size of the database segments for the years 1995 through 2005, the frequency of multiple segment records, and the increasing use of computers in a range of different crimes. We also illustrate the use of a join to extract relationships from the data.

Table 1 shows the number of NIBRS records in each of the non Batch NIBRS segments for the 10 year period beginning in 1995. The number of Administrative Segments records is the total number of incidents in the NBIRS database for the 1995–2005period and the table shows there are about 29 million incidents. The total number of records in each of the Offense, Property, Victim, Offender and Arrestee Segments is not much larger than the number of incidents in the Administrative Segment since few incidents involve multiple records. NIBRS make use of duplicate records to handle information uploaded in the year following an incident, which can affect incident counts. We explore this later.

Table 1. Sizes of NIBRS Segments
(1995 – 2005 data sets)

Segment Type	Record Counts	Attributes	Row Size (bytes)
Administrative	29,173,206	16	63
Offense	31,909,476	19	96
Property	33,348,435	25	77
Victim	31,741,442	55	142
Offender	32,940,666	11	54
Arrestee	8,015,646	19	96
Group B Arrest	9,865,679	21	74
Window Exceptional Clearance	11,502	26	97
Window Recovered Property	7,086	34	99
Window Arrestee	156,791	30	132

The Window segments allow ORIs to submit additional information for a Group A incident if information was not available for the initial submission. Submitting agencies use these segments to record exceptional events, for example, an ORI changes from a UCR summary submitter to NIBRS incident data submitter, a subsequent event causes the incident to be cleared, or an arrest is made at

after the initial submission. Table 6 shows a very small percentage of incidents involve Windows Segment submissions.

Not listed in the table are the segments Batch1, Batch 2 and Batch 3, which provide information on the ORI, population information, and submission information. Each of these segments contains 21,537 distinct ORI records for 2008, up from 18,643 in 1995. Although some information in Batch Segments is duplicated from year to year, we add these files each time we update the database since they include information on NIBRS agency participation trends.

Table 2 gives counts of incidents with single and multiple segments. The table shows about 5.8 million incidents have just one Arrestee Segment and about 726 thousand have two Arrestee Segments. For each incident in NIBRS, attributes in the Administrative Segment indicate the number of arrest, offender, offense and victim records associated with an incident. Notice that in the entire data collection (1995-2005) very few incidents have more than three segment records. For example, of the over 29 million incidents in the NIBRS 1995-2005 data collection, only about 688,000 (about 2.4%) have 3 or more offense records. Only about 235,000 incidents (about .8%) have four or more offense records. Here we only show results for up to 8 segments. The highest multiple segment count for an incident is 18 victim segments, which occurs in 163 incidents.
.

Table 2. Incidents with Multiple Records 1995-2005
(number of records per incident)

Segment	1	2	3	4	5	6	7	8
Arrestee	5,797,342	726,208	147,954	45,077	13,533	5,006	2,072	1,152
Offense	26,512,070	1,970,541	453,323	149,425	49,540	19,576	7,998	4,339
Offender	26,661,053	2,307,215	188,159	14,760	1,699	271	42	7
Victim	27,216,408	1,618,605	223,317	62,266	23,747	11,359	5,629	3,435

Collecting incident-based data on a national level and releasing it for research purposes presents a number of challenges. It may be necessary for reporting agencies to submit data to the FBI at different times if new data associated with an incident becomes available. In our implementation of NIBRS, the Administrative Segment contains two date attributes: the Incident Date and the NIBRS Release Date. The Incident Date is usually the date the incident occurred. If this date is unknown, NIBRS uses the date the incident was reported. The NIBRS Release Date is the date ICPSR assigns to the data set. Typically, the Incident year is the same as the NIBRS Release year. Table 3 shows the numbers and percentages of incidents where the NIBRS Release year differs from the NIBRS Incident year. The percentage of such incidents in the database is relatively small, 1.14% or less.

Table 3. NIBRS Released Year vs. Incident Year
(1995 – 2005 data sets)

NIBRS Released Year	Incidents where Released year is not the same as Incident year	Percentage of incidents where released year not the same as incident year	Total Release year Records
2005	52,351	1.13%	4,614,054
2004	46,490	1.14%	4,083,571
2003	39,856	1.10%	3,637,432
2002	36,941	1.07%	3,455,589
2001	200	*	3,232,281
2000	22,407	0.89%	2,841,523
1999	20,484	0.95%	2,157,326
1998	102	*	1,822,675
1997	56	*	1,426,978
1996	81	*	1,064,763
1995	168	*	837,014

*Data for these years reflect varying compliance standards in early NIBRS years

Currently, the FBI does not close out a NIBRS year until 14 months after the end of the year, in February of the following year. An incident that occurs and is reported to the FBI in December may not have complete information. The ORI may provide that information in the next year and it will appear in the NIBRS collection for the subsequent year. In such cases, NIBRS data will contain two identical incidents in subsequent NIBRS Release years. Care must be taken in counting incidents over multiple years since identical incident records can appear in different NIBRS release years. Sometimes the clearance date will indicate additional information for the incident that became available later. Often an arrestee record date will indicate that the arrest was made in a subsequent year. The arrestee date is frequently updated in the first year the incident appears. As one might expect, incidents that occur late in the year are far more likely to result in a duplicate record in a subsequent year.

The relational database facilitates multiyear analysis. Table 4 shows the percentages of offense where the offender is suspected of using a computer. The table provides a snapshot of NIBRS data at five year intervals. As expected, wire fraud and pornography/obscene materials offenses are far more likely to involve a computer than other offenses. Although the reported computer use for most crime categories is small, most categories show considerable increases from 1995 to 2005. Certain categories such as betting/wagering, drug/narcotics violations, extortion blackmail, and intimidation show a leveling off or a slight decline from 2000 to 2005.

Table 4. Offender Suspected of Using a Computer
(Percentages of Offenses)

Offense	1995	2000	2005
Betting/Wagering	0.00	0.63	0.30
Confidence Game	0.22	0.68	1.53
Counterfeiting/Forgery	0.04	0.78	1.11
Drug/Narcotic Violations	0.06	0.15	0.08
Embezzlement	0.52	0.47	0.90
Extortion/Blackmail	0.00	0.99	0.79
Gambling Equipment Violations	0.00	0.47	0.86
Impersonation	0.00	0.97	1.51
Intimidation	0.05	0.31	0.29
Gambling	0.00	0.49	0.33
Pornography/Obscene Material	0.00	10.42	16.43
Robbery	0.02	0.06	0.02
Statutory Rape	0.00	0.17	0.20
Teller Machine Fraud	0.37	1.83	2.39
Wire Fraud	3.33	13.17	18.81

This type of NIBRS data raises several questions. First, the unit of analysis here is the offense and all data are derived from the offense segment. Second, each offense record contains three attributes, Offender Suspected of Using 1, 2 or 3. Each attribute can have one of the responses: alcohol, drugs, computer equipment or not applicable. The SQL statement that generated the above data looked for the code for "computer use" in each of these fields and assumed the data were consistent, i.e., no two fields indicated that computer equipment was involved. In multiyear analysis, the increase in computer use may simply be due to better gathering and reporting systems in use by NIBRS reporting agencies, which increase the likelihood that data will be provided for this attribute. The consistency of results over different crime categories, however, indicates the results are not just due to the way the data were collected. The effect of non response for particular data elements in NIBRS is not known and requires further study.

As noted our implementation follows NIBRS and simply creates separate attributes for multiple response data elements. Listing these multi response data elements as separate attributes rather than implementing them in a separate table simplifies the database and facilitates processing at the cost of possibly slightly more storage. The victim segment has the most multiple response data elements. For example, NIBRS will record up to five possible injuries suffered by a victim. Each is listed as a separate attribute in the victim segment, injury 1,2,3,4 or 5. A study of NIBRS victim records shows only about 21% of the victim records have injury 1 populated with a code; 50% of those records (or 11% of all victim records) contain the injury type "none" indicating that the victim did not suffer an injury. The data indicate injury attributes 2 and higher are seldom used[7]. The Victim Segment also records up to 10 different offenses associated with the victim. Again, each is a separate attribute in the table. The Victim Segment also lists up to 10 different offenders and their relationship to the victim. About 25% of victim records list one or more offender/victim relationships, about 5% list two or more and only about 1% list three or more relationships.

[7] Of the 48.2 million victim records in 1995-2008 NIBRS collection, about 10.1 million have a code listed for injury 1 and are not null. For Injury 2 less than 70 thousand (less than .2% of victim records) have a code listed. Use of the remaining injury attributes is negligible.

Studies that require NIBRS data in different segments have long posed difficulties for data analysts. In a relational data base, the join is the standard tool to merge data in two or more different segments or tables. Even though this might appear computationally intensive, the fact that a NIBRS incident is likely to involve only a few offender, victim or offense records limits the computational costs. On most modern equipment, such computations are not that time consuming. On our systems for the NIBRS collection from 1995-2008 a join of the victim and offender tables , two segments frequently used together, takes about 5 minutes and produces a table with approximately 57.5 million rows[8]. For the 1995-2008 collection, the victim and offender tables have about 48 million and 50 million rows, respectively. Queries in the table produced by the victim-offender join typically require less the one minute.

One study of NIBRS data often of interest is the relationship between the age of the victim and offender. Here we show a join of the victim and offender tables that provides a list of incidents sorted by ORI code, incident number, offender sequence number and victim sequence number. A sample of the resulting table appears as Table 5. The table highlights an incident where there are multiple victims and offenders. For the incident beginning at line 3, we see there are 4 offenders and 2 victims. The table clearly shows the relationships between the ages of victims and offenders. The data in the complete table show a strong correlation between the age of the victim and offender. In addition, there is a strong correlation between incidents that have multiple victims and offenders and the ages of those involved.

Table 5. Victim/Offender Join

	ORI Code	Incident Number	Offender Sequence No.	Offender Age	Victim Sequence No.	Victim Age	Incident Date
1	CO0030400	CI0BRFRH-2 N	1	23	3	33	09-Nov-00
2	DE0020300	LT01KETVV0 N	0	00	1	39	16-DEC-02
3	DE0020600	LI01KVBRTU N	1	11	1	09	06-OCT-0
4	DE0020600	LI01KVBRTU N	1	11	2	08	06-OCT-0
5	DE0020600	LI01KVBRTU N	2	10	1	09	06-OCT-0
6	DE0020600	LI01KVBRTU N	2	10	2	08	06-OCT-0
7	DE0020600	LI01KVBRTU N	3	10	1	09	06-OCT-0
8	DE0020600	LI01KVBRTU N	3	10	2	08	06-OCT-0
9	DE0020600	LI01KVBRTU N	4	12	1	09	06-OCT-0
10	DE0020600	LI01KVBRTU N	4	12	2	08	06-OCT-0
11	IA0820200	7Z1C7REMQ-F	1	40	1	41	24-JAN-02

In computing the join we included records where the offender's sequence number might be 0. (See row 2 in Table 5.) NIBRS uses this code value to indicate that nothing is known about the offender, possibly not even the number of offenders. Ages are stored as character strings in NIBRS, and the age attribute normally would contain an age, as is the case with most entries in the age column in the table.

[8] Our relational database implemention of NIBRS uses an Oracle 11G Database running on an HP Proliant ML 370G5 Server with 8 Giga Bytes of RAM. Data partitioning was note used.

A code value of '00' indicates that the age was either unknown or left blank. Of the 50,036,343 Offender Segment records in the 1995–2008 NIBRS collection, 12,853,346 (about25%) have age value of '00' and about 15,841,074 (about 32%) have a sequence number of 0. Thus, missing data in NIBRS may impact the conclusions we reach using NIBRS data for victim-offender age studies.

We have notice a number of other NIBRS incident variables are missing or frequently not reported. For example, NIBRS gives investigators the opportunity to record the circumstances surrounding a criminal activity. The circumstances are a multi response variable in the Victim Segment and up to two circumstances may be recorded. NIBRS provides codes for the following: argument, assault on law enforcement, drug dealing, gang activity, child playing with weapon. (A complete list appears in the codebook). In only 4 to 5% of victim records is there at least one code that indicates the circumstances surrounding an incident. Moreover, this frequency is remarkably consistent throughout the 1995-2008 collection.

As illustrated, having NIBRS in a relational database facilitates a variety of studies. Moreover, one can employ the power of the join, a feature in the SQL Select statement, to extract relationships. In addition, modern databases employ various techniques such as partitioning and indexing to optimize data processing. Here we have shown that despite rather large tables, a modern database can take advantage of the NIBRS structure to perform an SQL join of the largest NIBRS tables on fairly modest computing equipment at reasonable cost. Finally, databases usually provide checks for data referential integrity which ensures our database has all NIBRS codes needed when new data are download from NACJD.

6. Precomputed views of NIBRS data and interactive analysis capabilities

In this section we use our relational database and modern spreadsheets capabilities to enable analysts to create views of NIBRS data. The basic idea is to let a database expert use SQL commands to generate the raw data needed for a range of studies, and then let the analyst import the data into a spreadsheet and develop views of the data needed for his or her particular study[9]. Here we extract NIBRS data that allows researchers to study offender demographics and the use of a computer in various offenses. Using modern spreadsheet features such as Pivot Tables found in Microsoft Excel, analysts can perform analyses similar to those done with data cubes, for example, aggregation and disaggregation according to different demographic features. In addition, modern spreadsheets provide a wide range of statistical and report capabilities that can be applied to the data.

We assumed the analyst was interested in studying offenders who committed various types of offenses where a computer played a role. We were interested in providing capabilities that would allow analysts to aggregate counts of offenders according to a specific offense and demographic features such as race, age and gender. Thus, the data would allow analysts to determine if an offender in a particular age group might be more likely to use a computer in a given type of crime. In addition to selecting the offenses with the pivot table feature, users also can select the years for which counts are presented. In order to study age, we divided ages up into the following 10 year categories: child (0-10); teen (11-20); young adult (21-20); three middle age categories 31-40, 41-50, 51-60; and two senior categories 61-70, and 71-80. With the pivot table feature, users have the option of combining these categories and simply studying the offender counts aggregated according to race and gender.

[9] Of course the data could also be imported into a statistical package such as SAS or SPSS.

Generating the raw data for the spreadsheet pivot table required linking data elements in the Offense and Offender NIBRS segment tables. In NIBRS, the multi response data element that indicates computer involvement, "Offender Suspected of Using (1-3)," is in the Offense Segment. When provided, the possible responses for the Offender Suspected of Using attributes are a combination of one or more of the following: drugs, alcohol, computer equipment, or not applicable. The code for computer equipment may appear in any one of the three attributes which complicates processing. The type of crime is indicated by the UCR offense code attribute in the Offense Segment. The offender demographic attributes, age, gender and race, are in the Offender Segment. If an offense is linked to multiple offenders, NIBRS does not distinguish if a particular offender made use of a computer and another did not. Thus, if four offenders are linked to an incident involving one offense, we would count four offenders as having made use of a computer in the crime. In this study offenders are counted for each offense. If an offender is involved in two offenses, the offender will appear in the counts for each of those offenses.

An SQL join of the Offender and Offense Segment tables produces the intermediate table that contains the data imported into the spreadsheet pivot table. (See Appendix III) Attributes in the intermediate table, which are derived from the original segment tables, include the NIBRS Released Year; the ORI and Incident Codes; a description of the UCR offense; and the race, age group and gender of the offender. The age groups were calculated as part of the SQL join. We restricted the list of records in the intermediate table to include only those incidents where the released year matches the incident year to avoid record duplication. (We have already discussed the problem of duplication in multiyear NIBRS studies.) Once the intermediate table (or raw data) is imported into the spreadsheet many different studies are possible. The three tables presented in this section were derived from one pivot table and the raw data generated by the join. The spreadsheet pivot table feature provides clickable buttons that enable aggregation and disaggregation of various demographic features, inclusion or exclusion of features, selection of offenses, and selection of years.

Table 6 shows offender counts for pornography offenses in which the offender is suspected of using a computer. The offender counts are aggregated by age groups. In this study the analyst has chosen to display counts for the years 1995 to 2005. The youngest and oldest age groups are not shown since in each case only one offender was listed for the entire period.

Table 6 Offender Counts (Offender suspected of using a computer)

Offense	Age	'95	'96	'97	'98	'99	'00	'01	'02	'03	'04	'05	Total
Pornography Obscene Material	11-20			1	3	12	20	25	22	43	40	60	226
	21-30		1		7	7	21	20	30	39	79	74	278
	31-40		1	2	10	6	10	29	40	42	62	70	272
	41-50		1	3	6	10	17	22	17	38	59	52	225
	51-60				2	1	4	12	17	37	30	33	136
	Totals		3	6	28	36	72	108	126	199	270	289	1,137

Using the pivot table field list feature, we can deselect age and view the offender counts aggregated by race. Table 7 shows offenders aggregated by race while age is suppressed. The offender counts listed for various NIBRS racial categories, Black, White, Unknown, Native American or Blank. Columns again show how these values change over time. Here again only one offense is shown.

Table 7 Offender Counts (Offender suspected of using a computer)													
Offense	Race	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Total
Pornography Obscene Material	A					1	2	4	3	3	5	5	25
	B		5	2	120	67	98	228	115	166	188	236	1,125
	I				1	2	1	2	2			2	9
	U	3	1	9	59	101	157	376	445	245	208	185	1,858
	W	2	14	49	142	141	262	556	400	394	417	704	1,942
	Blank									137	216	381	1,114

The pivot feature allows us to do various studies on the raw data produced by a join of the offender and offense segments. Table 8 shows offender counts for two selected offenses, embezzlement and wire fraud. Here we aggregate offenders according to age and gender. We also choose to display only the years 2000 to 2005.

Table 8 Offender Counts (Offender suspected of using a computer)										
Count of offenders: Aggregated by Offense, Age and Gender			Year							
Offense Description	Age Group	Gender	2000	2001	2002	2003	2004	2005	Grand Total
Embezzlement	11 – 20	F	11	7	5	8	8	14	53
		M	4	5	6	8	6	7	36
	20 – 30	F	17	18	19	22	20	29	125
		M	11	13	14	14	12	23	87
	31 – 40	F	9	9	18	20	13	31	100
		M	8	9	12	7	12	13	61
	41 – 50	F	5	7	7	6	8	21	54
		M	3	8	4	4	4	10	33
	51 – 60	F		2	4	1	1	4	12
		M	1	4			2	3	10
Wire Fraud	11 – 20	F	1	3	3	4	2	2	15
		M	9	9	9	13	13	12	65
	20 – 30	F	1	6	3	6	14	16	46
		M	7	12	18	22	27	22	108
	31 – 40	F		3	2	8	9	8	30
		M	4	8	11	12	13	21	69
	41 – 50	F	1	5	3	3	3	6	21
		M	4	2	2	8	4	5	25
	51 – 60	F		1	3	1	2	4	11
		M	2	2	2	1	1	6	14

The preceding tables only show counts of offenders for several different offenses. These are presented simply to demonstrate some of the capabilities available in most modern spreadsheets. Percentages could also easily be computed.

The numbers presented indicate that the Offender Suspected of Using Attribute frequently is left blank. Other authors have noted that certain incident level details often are missing in NIBRS data. Addington (2008), for example, examined the frequency with which the Circumstance attribute in the Victim Segment is left out in homicide cases. (About 35% for NIBRS compared with 30% for the FBI's Supplementary Homicide Reports.) Non response for various NIBRS data elements and their effect on analysis certainly warrants further study.

This section illustrates one of the main values of a relational database implementation of NBIRS – it facilitates the preparation of NIBRS data to support a variety of specialized studies. We are currently working with criminal justice researchers to provide raw data to support studies in several areas. We plan to make a library of such collections available for download; thereby making otherwise often difficult to extract NIBRS data available to the criminal justice research community.

7. Concluding Remarks

In this paper we have described some of the facilities available for analyzing NIBRS data including our own relational database implementation of NIBRS. The relational database implementation offers many advantages; however, the analyst needs to work with a data processing specialist. With a relational database implementation, data processing specialists can develop views of NIBRS data suitable for particular studies. Researchers can then use powerful tools such as spreadsheet pivot tables to selectively analyze the data. Finally, a relational database allows us to study NIBRS data quality and take full advantage of the rich collection of incident-based crime data in the NIBRS.

Acknowledgement

The authors gratefully acknowledge Boris Bondarenko and Raul Cabrera who helped develop the initial implementation of the NIBRS relational database. The authors also thank the FBI Criminal Justice Information Services, the ICPSR-NACJD, and the NYS Justice Services Division for their helpful advice and willingness to take the time to answer our many questions. Finally, the authors thank Tom Holt of The Michigan State School of Criminal Justice for suggesting the study topic in section 6.

References

Addington, L., (2004). The Effect of NIBRS Reporting on Item Missing Data in Murder Cases. Homicide Studies, 8, 193-213.

Addington, L. (2008). Assessing the Extent of Nonresponse Bias on NIBRS Estimates of Violent Crime. Journal of Contemporary Criminal Justice, 24, 32-49.

Akiyam, Y. (October 1997). Cross Tabulations and Units of Count with NIBRS Data Elements. Federal Bureau of Investigation, Criminal Justice Information Services Divsion.

Akiyama, Y., & Nolan, J. (1999). Methods for Understanding and Analyzing NIBRS Data. Journal of Quantitiative Methods in Criminology , 15, 225-238.

Chilton, R., & Regoeczi, W. (2005). The Importance of being specific: Offense, Victim, Offender, and Arrest Rates in NIBRS. presented at the annual meeting of the American Sociological Association. Philadelphia, PA: All Academic, Inc.

Dunn, C. S., & Zelenock, T. J. (1999). NIBRS Data Available for Secondary Analysis. Journal of Quantitative Criminology , 15 (2), 239-248.

Edgar, P., Cohen, J., & Porter, J. (2004). Web-Based Data Resources at the NCOVR Data Center. Carnegie Mellon University , National Consortium on Violence Research, Pittsburgh .

Federal Bureau of Investigation. (2009, January 14). About the Uniform Crime Reporting Program. Retrieved Februray 10, 2011, from Uniform Crime Reporting Statistics: http://www.ucrdatatool.gov/abouttheucr.cfm

Federal Bureau of Investigation. (2008). National Incident-Based Reporting System, 2008: Codebook. United States Department of Justice, FBI. Ann Arbor, Michigan: Inter-University Consortium for Political Science and Social Research.

JRSA. (2008, February). Status of NIBRS in the States. Retrieved January 2011, from IBR Resource Center: http://www.jrsa.org/ibrrc/background-status/nibrs_states.shtml

Maxfield, M. (1999). The National Incident-Based Reporting System: Reserach and Policy Applications. Journal of Quantitative Criminology , 119-149.

NACJD. (2010). National Incident-Based Reporting System Resource Guide. Retrieved February 10, 2011, from National Archive of Criminal Justice Data: http://www.icpsr.umich.edu/icpsrweb/NACJD/NIBRS/

NACJD. (2008). National Incident-Based Reporting System, 2008 : Extract Files. Ann Arbor, Michigan: Inter-university Consortium for Political and Social Science Research .

NACJD. (2010). NIBRS Data and File Specification. Retrieved March 15, 2011, from National Archive of Criminal Justice Data: http://www.icpsr.umich.edu/icpsrweb/NACJD/NIBRS/specs.jsp

NAJCD. (2010). Retrieved January 3, 2011, from Data Resource Guides: http://www.icpsr.umich.edu/icpsrweb/NACJD/

NCOVR. (n.d.). Access NCOVR Data Sets. Retrieved Novmeber 10, 2011, from National Consortium on Violence Research: http://www.ncovr.heinz.cmu.edu//Docs/Direct_access_data_cubes_and_discoverer.htm#

Edgar, P., Cohen, J., & Porter, J. (2004). Web-Based Data Resources at the NCOVR Data Center. Carnegie Mellon University , National Consortium on Violence Research, Pittsburgh .

NCJJ. (2010, September 23). Easy Access to NIBRS: Victims of Domestic Violence 2008. Retrieved May 10, 2011, from http://www.ojjdp.gov/ojstatbb/ezanibrsdv/

Roberts, A. (2009). Contributions of the National Incident-Based Reporting System to Substantive Knowledge in Criminology: A Review. Sociology Compass , 3 (3), 433-458.

SDA. (2011, March 29). SDA:Survey Documenation and Analysis System. Retrieved April 10, 2011, from http://sda.berkeley.edu/document.htm

United States Department of Justice, Federal Bureau of Investigation. (2010, September). *Crime in the United States 2009*. Retrieved March 2011, from http://www2.fbi.gov/ucr/cius2009/about/index.html

Appendix I – NIBRS Adoption

NIBRS adoption for reporting crime data to the FBI has been progressing albeit slowly. Since NIBRS is a voluntary program, adoption by law enforcement agencies, especially in larger municipalities, has been slow. According to the Justice Research Association (JRSA, 2008), as of Feb. 2008 about 37% of law enforcement agencies submitted crime data through the NIBRS program, which accounts for about 25% of the US population and is up from about 14% of the population in 2000. As of 2008 the FBI had certified 31 states to submit crime data through the NIBRS. Currently, the largest reporting municipality is Fairfax County Virginia, which has a population just under one million. The FBI reports that for crime data submitted for 2009, about 44% of the reporting agencies submitted their data through the NIBRS program. These submissions accounted for 26% of the crime data collected and covered about 28% of the population (United States Department of Justice, Federal Bureau of Investigation, 2010).

Despite programs to promote NIBRS reporting, municipalities have been reluctant to devote resources to developing systems that would not directly benefit operations. NIBRS after all is for crime analysis and does not provide information such as names and locations needed for an operational system. However, reporting agencies and states increasingly are using operational systems based on NIBRS standards. These systems facilitate NIBRS reporting as a byproduct[10]. Based on the NIBRS standards, these systems provide information needed for use in operations, tools to glean intelligence, and various mapping tools to aid analysis.

NIBRS provides a standard for sharing crime information. Although separate programs, NIBRS provides the structure for data in submitted to the FBI's National Law Enforcement Data Exchange Program (NDEX). In fact, NIBRS data can be submitted to the FBI through NDEX (FBI web site and Police Chief Paper) but NIBRS submission for agencies that use NDEX is not required[11].

NIBRS was designed to replace the UCR summary reporting system. NIBRS participation is increasing and agencies that participate are collecting more complete NIBRS incident data, especially as more adopt operating systems based on NIBRS. As noted we have seen a 51% increase in the number of incidents in the NIBRS data just in the years 2005-2008. Also, additional NIBRS participation is expected to increase as reporting agencies replace or upgrade outdated operational record management systems with systems that are NIBRS compliant.

[10] http://www.zuerchertech.com/products/state_agency_solutions/nibrs, http://ucr.psp.state.pa.us/ucr/News/ComNewsUI.asp?ReqPage=AboutUCRToday.htm http://ocjs.ohio.gov/Oibrs_Web/Data_Specs_2009.pdf

[11] http://ucr.psp.state.pa.us/ucr/News/ComNewsUI.asp?ReqPage=AboutUCRToday.htm

Appendix II - Group A Segment Descriptions

ADMINISTRATIVE SEGMENT
NIBRS_Release_Date State_Code ORI_Code Inicident_Number INCIDENT_DATE
INCIDENT_DATE_HOUR TOTAL_OFFENSE_SEGMENTS TOTAL_VICTIM_SEGMENTS
TOTAL_OFFENDER_SEGMENTS TOTAL_ARRESTEE_SEGMENTS CITY_SUBMISSION
CLEARED_EXCEPTIONALLY EXCEPTIONAL_CLEARANCE_DATE NUMBER_OF_RECORDS_PER_ORI

ARRESTEE SEGMENT
NIBRS_RELEASE_ DATE, STATE_CODE, ORI_CODE INCIDENT_NUMBER INCIDENT_DATE
ARRESTEE_SEQUENCE_NUMBER ARREST_TRANSACTION_NUMBER ARREST_DATE TYPE_OF_ARREST
MULTI_ARRESTEE_SEG_INDICATOR UCR_ARREST_OFFENSE_CODE ARRESTEE_WAS_ARMED_WITH (1-2) AGE_OF_ARRESTEE SEX_OF_ARRESTEE RACE_OF_ARRESTEE ETHNICITY_OF_ARRESTEE
RESIDENT_STATUS_OF_ARRESTEE
DISPOSITION_ARRESTEE_UNDER18 NUMBER_OF_RECORDS_PER_ORI

OFFENDER SEGMENT
NIBRS_RELEASE_ DATE STATE_CODE ORI_CODE INCIDENT_NUMBER INCIDENT_DATE
OFFENDER_SEQUENCE_NUMBER AGE_OF_OFFENDER SEX_OF_OFFENDER
RACE_OF_OFFENDER NUMBER_OF_RECORDS_PER_ORI

PROPERTY
SEGMENT, NIBRS_RELEASE_ DATE, STATE_CODE, ORI_CODE, INCIDENT_CODE INCIDENT_DATE
TYPE_PROPERTY_LOSS PROPERTY_DESCRIPTION VALUE_OF_PROPERTY DATE_RECOVERED
NUMBER_STOLEN_VEHICLES, NUMBER_RECOVERED_VEHICLES
SUSPECTED_DRUG_TYPE (1-3) ESTIMATED_QUANTITY (1-3) ESTIMATES_QUANTITY_FRACTIONAL (1-3)
TYPE_MEASUREMENT (1-3)
NUMBER_OF_RECORDS_PER_ORI

OFFENSE SEGMENT
NIBRS_RELEASE_ DATE STATE_CODE ORI_CODE INCIDENT_NUMBER INCIDENT_DATE
UCR_OFFENSE_CODE OFFENSE_ATTEMPTED_COMPLETED OFFENDER_SUSPECTED_USING (1-3)
LOCATION_TYPE NUMBER_OF_PREMESISES_ENTERED METHOD_OF_ENTRY
TYPE_CRIMINAL_ACTIVITY (1-3) TYPE_WEAPON_FORCE_INVOLVED (1-3)
BIAS_MOTIVATION NUMBER_OF_RECORDS_PER_ORI

VICTIM SEGMENT
NIBRS_RELEASE_ DATE STATE_CODE, ORI_CODE INCIDENT_NUMBER INCIDENT_DATE
VICTIM_SEQUENCE_NUMBER UCR_OFFENSE_CODE (1-10) TYPE_OF_VICTIM AGE_OF_VICTIM
SEX_OF_VICTIM RACE_OF_VICTIM ETHNICITY_OF_VICTIM RESIDENT_STATUS_OF_VICTIM
CIRCIMSTANCE (1-2) ADDITIONAL_JUSTIFIABLEHOMICIDE INJURY (1-4)
OFFENDER_NUMBER_RELATED (1-10) RELATIONSHIP_VICTIMOFFENDER (1-10)
OTHER_ORI_JURISDICTION TYPE_ACTIVITY TYPE_ASSIGNMENT NUMBER_OF_RECORDS_PER_ORI

```
rem ============================================
rem Date : 10/31/2010
rem SQL Code to list the offenders using computers in all NIBRS offenses.
rem The age, gender, and race of each offender is listed.
rem Also listed are the UCR offense code description and the NIBRS release
rem year for the incident.
rem Subtotals are listed for incidents involving multiple offenders.
rem Table prepared for import into an Excel Spreadsheet.
rem ============================================


SELECT nibrs_released,
       state_code,
       ori,
       description,
       age_of_offender_group,
       sex_of_offender,
       race_of_offender,
       number_of_incidents
FROM   (SELECT nibrs_released,
               state_code,
               ori,
               ucr_offense_code,
               age_of_offender_group,
               sex_of_offender,
               race_of_offender,
               COUNT(*) AS number_of_incidents
        FROM   (SELECT a.nibrs_released,
                       a.state_code,
                       a.ori,
                       a.ucr_offense_code,
                       ( CASE
                           WHEN age_of_offender > '00'
                                AND age_of_offender <= '10' THEN 'CH'
                           WHEN age_of_offender > '10'
                                AND age_of_offender <= '20' THEN 'TE'
                           WHEN age_of_offender > '20'
                                AND age_of_offender <= '30' THEN 'YO'
                           WHEN age_of_offender > '30'
                                AND age_of_offender <= '40' THEN 'M1'
                           WHEN age_of_offender > '40'
                                AND age_of_offender <= '50' THEN 'M2'
                           WHEN age_of_offender > '50'
                                AND age_of_offender <= '60' THEN 'M3'
                           WHEN age_of_offender > '60'
                                AND age_of_offender <= '70' THEN 'S1'
                           WHEN age_of_offender > '70'
                                AND age_of_offender <= '99' THEN 'S2'
                           ELSE age_of_offender
                         END ) AS age_of_offender_group,
                       b.sex_of_offender,
                       b.race_of_offender
                FROM   nibrs.offense a,
                       nibrs.offender b
                WHERE  a.nibrs_released = b.nibrs_released
```

```
                AND a.state_code = b.state_code
                AND a.ori = b.ori
                AND a.incident = b.incident
                AND a.incident_date = b.incident_date
                AND a.nibrs_released=Extract(YEARFROM a.incident_date)
                AND b.nibrs_released=Extract(YEARFROM b.incident_date)
                AND ( offender_suspected_using1 = 'C'
                      OR offender_suspected_using2 = 'C'
                      OR offender_suspected_using3 = 'C' ))
 GROUP  BY nibrs_released,
           state_code,
           ori,
           ucr_offense_code,
           age_of_offender_group,
           sex_of_offender,
           race_of_offender) a
left outer join ucr_offense_code b
  ON a.ucr_offense_code = b.code;
```