LA CRIMES 2012 - 2015

CS 434 Project 2

Joshua Haupt, Charles Stewart Helf
jhaupt@siue.edu, malkav.amonra@gmail.com

DBMS: MySQL / MariaDB on Linux

Database Schema:

Charles, please look into adding constraints

PoliceDepartment(

       PrecinctNumber: INT(incr),

       Jurisdiction: VARCHAR(35)

)

PoliceOfficer(

       DateGraduated: DATE,

       BadgeNumber: INT,

       Name: VARCHAR(50)

)

CrimeReport(

       ReportNumber: INT(incr),

       DateFiled: DATE,

       Description: VARCHAR(250),

)

StatusUpdate(

       ReportNumber: INT(incr),

       RevisionNumber: INT(incr),

       Status: VARCHAR(1),

       Date: DATE

)

CriminalIncident(

       IncidentNumber: INT(incr),

       TimeOccurred: INT,

       DateOccurred: DATE,

Address: VARCHAR(50)

)

Statute(

CodeDesignation: VARCHAR(12),

ElementOfCrime: VARCHAR

)

Relations:

Relation 1: Statute(Code Designation, Elements of Crime)

Code Designation → Elements of Crime

Relation 2: Criminal Incident(Incident Number, Time, Date, Address)

Incident Number → Time, Date, Address

Relation 3: Defined By[**Statute - Criminal Incident**] (Code Designation, Incident Number)

Incident Number → Code Designation

Relation 4: Crime Report(Report Number, Date Filed, Jurisdiction)

Report Number → Date Filed, Description

Relation 5: Reported Through[**Criminal Incident - Crime Report**] (Incident Number, Report Number)

Incident Number → Report Number

Report Number → Incident Number

Relation 6: Status Update(Report Number, Revision Number, Date, Status)

Report Number, Revision Number → Date, Status

Report Number, Date → Revision Number, Status

Relation 7: Given A[**Crime Report** - **Status Update**] (Report Number)

None

Relation 8: Police Officer(Badge Number, Name, Graduation Date)

Badge Number, Graduation Date → Name

Relation 9: Filed By[**Crime Report - Police Officer**] (Graduation Date, Badge Number, Report Number)

Graduation Date, Badge Number → Report Number

Relation 10: Police Department(Precinct Number, Jurisdiction)

Precinct Number → Jurisdiction

Relation 11: Member Of[**Police Officer** - **Police Department**] (Precinct Number, Graduation Date, Badge Number)

Graduation Date, Badge Number → Precinct Number

An Overview of the Schema

We were unable to find any non-BCNF relations in our schema. Thus, the current layout of our data and the relations described therein should, at least superficially, avoid producing redundancies in our database. This is in part due to the rather succinct nature of the data set we're using, but it would be remiss to ignore the role research into the real world played in abetting this characteristic. We paid attention to how the data interacts in real life, and it seems to have paid off. Tables should contain only exactly what they need to describe their intended relations and should join with others in a very straightforward manner. That said, two of our six entities—and, by extension, the relationships they have with others—are weak. It's unusual. It also necessitates duplicating attributes. In effect, we end up relying on duplicating some data to ensure that we can properly link it to other data. This is probably inescapable, but given such a small amount of variety in the data set, avoiding any data duplication at all would have been preferable.

Since criminal incidents and crime reports are at the figurative (and literal) center of the diagrams, we clearly expect that most of this databases queries will revolve around patterns in crime reporting with regard to the actual criminal incidents, their status updates, the police who respond to incidents and conduct reports, and so on. Curious users may want to look at less intuitive patterns, such as for which law violations a particular police department makes more frequent arrests, or which officers are more (or less) likely to make frequent updates to crime reports for particular crimes. This data could be used to answer more interesting questions, such as, "What crimes does the North Hollywood Area department not take very seriously?" or, "Does Officer Jared Lyndon slack off when it comes to investigating vandalism incidents?"

Relational Schema

STATUTE

Attribute	Code Designation	Elements of Crime
Data Type	Varchar(12)	Int**

Relation 1: {Code Designation, Elements of Crime}

Code Designation → Elements of Crime

Defined-By

Attribute	Incident Number	Code Designation
Data Type	Int (incrementing)	Varchar (12)

Relation 2: {Code Designation, Incident Number}

Incident Number → Code Designation

CRIMINAL INCIDENT

Attribute	Incident Number	Time Occurred	Date Occurred	Address
Data Type	Int (incrementing)	Int (4-digit time)****	Date	Varchar(50)

Relation 3: {Incident Number, Time Occurred, Date Occurred, Address}

Incident Number → Time Occurred, Date Occurred, Address

Reported-Through

Attribute	Report Number	Incident Number
Data Type	Int (incrementing)	Int (incrementing)

Relation 4: {Report Number, Incident Number}

Report Number → Incident Number
Incident Number → Report Number

CRIME REPORT

Attribute	Report Number	Date Filed	Description
Data Type	Int (incrementing)	Date	Varchar(200)

Relation 5: {Report Number, Date Filed, Description}

Report Number → Date Filed, Description

Given-A

Attribute	Report Number	Report Number
Data Type	Int (incrementing)	Int (incrementing)

Relation 6: {Report Number}

STATUS UPDATE

Attribute	Report Number	Revision Number	Date Revised	Status

Data Type	Int (incrementing)	Int (incrementing)	Date	Int***

Relation 7: {Report Number, Revision Number, Date Revised, Status}

Report Number, Revision Number → Date Revised, Status
Report Number, Date Revised → Revision Number, Status

Filed-By

Attribute	Date Graduated	Badge Number	Report Number
Data Type	Date	Int	Int (incrementing)

Relation 8: {Date Graduated, Badge Number, Report Number}

Date Graduated, Badge Number → Report Number

POLICE OFFICER

Attribute	Date Graduated	Badge Number	Last Name	First Name
Data Type	Date	Int	Varchar(25)	Varchar(25)

Relation 9: {Date Graduated, Badge Number, Last Name, First Name}

Date Graduated, Badge Number → Last Name, First Name

Member-Of

Attribute	Date Graduated	Badge Number	Precinct Number
Data Type	Date	Int	Int (tiny)

Relation 10: {Date Graduated, Badge Number, Precinct Number}

Date Graduated, Badge Number → Precinct Number

POLICE DEPARTMENT

Attribute	Precinct Number	Jurisdiction
Data Type	Int (tiny)	Varchar(50)

Relation 11: {Precinct Number, Jurisdiction}

Precinct Number → Jurisdiction

Notes:

** Since "Elements of a Crime" is essentially a pooled checklist attribute--that is, it's a series of true/false flags for a pool of common items--it's suggested that it be modeled this way in the data somehow. This way, it will reduce the amount of disk space immensely while simultaneously providing richness of information.

For instance, Aggravated Battery and Murder both contain the crime elements "interpersonal contact" and "injury", but Murder would also contain the "resulting in death" element, whereas Aggravated Battery would not. Similarly, Robbery and Grand Burglary would both contain the crime elements "possession" and "property belonging to another", but whereas Robbery would also contain "interpersonal contact" and "threat of use of force" elements, Burglary would instead contain the "breaking into premises" element.

*** Since this is a multiple choice-style data element, a very small integer or set of booleans should suffice. This can be implemented in an almost identical manner to the above note.

****Since the times are only reported to the minute and not to the second, the TIME data type would not be efficient in our case. Therefore, we chose to go with the INT data type as we only need to record time down to the minute in a 24-hour format.