Using an Insurance Company Database Example to Teach Database Design and Data Mining Principles

Julia Bao (U. of Phoenix)
juliabao@email.phoenix.edu

University of Phoenix®

and
Chris Helwig (EMCC)
WORLDCOMP'13

ABSTRACT

This project involves using an insurance company's data to illustrate database design principles. Students are told to suppose they are designing a new database system for an insurance company. The insurance company provides auto, homeowners, renters, and life insurance policies to its customers. Students must identify what kinds of data to place in database tables, what data types to use, what primary and secondary keys to use, and what relationships to enforce between database tables. To complete this task, it is necessary to understand the insurance company's business model, as well as to consider how the database might later be used as an analytical tool and as a data repository.

General Terms

Computer Science Education, Database Design, Microsoft Access.

Keywords

Insurance Company Database, Database Schema.

1. The database could include the following tables

1. Customer Info, with fields for ID, Last Name, First Name, Address, City, State, Zip, Phone, Credit Card Number, Exp Date, Multi Policy Discount, DOB.
2. Auto, with fields for Auto Policy Number, ID, Make, Model, Year, Liability Amount, UM, UIM, Med Pay, Premium Amount, Collision Damage Amount, Named Insured, Additional Driver.
3. Homeowners, with fields for Homeowners Policy Number, ID, Liability Amount, Property Damage Amount, Premium.
4. Renters, with fields for Renters Policy Number, ID, Liability Amount, Property Damage Amount, Premium.
5. Life Ins, with fields for Life Ins Policy Number, ID, Premium amount, Benefit Amount, Beneficiary.

The Customer Info table would contain basic customer contact, biographical information, and payment information. The Auto table would contain information related to a customer's automobile policy. UM refers to uninsured motorist coverage and UIM refers to under-insured motorist coverage. This table contains as a foreign key the Customer Info table ID field. The Homeowners table contains information related to Homeowners insurance coverage, the Renters table contains information related to Renters insurance, and the Life Ins table contains information related to Life insurance.

Also needed would be transaction tables for transaction related data:

1. Payments, with fields for Payment Transaction ID, ID, Due Date, Amount Due, Amount Paid, Paid on Time.
2. Claims, with fields for Claim Transaction ID, ID, Amount of Claim, Claim Approved.

The database schema should avoid data redundancy, maintain data integrity, and organize data in a logical manner.
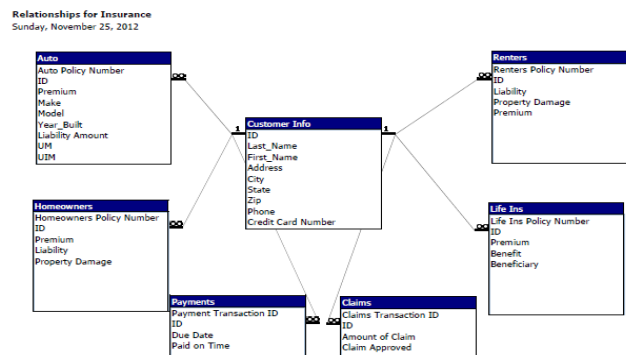

Figure 1. Screen shots of what this database would look like in Microsoft Access.

2. Further Database Details

This relationship report above was generated in Access under Database Tools | Relationships and Relationship Tools | Design | Relationship Report. A one to many relationship with referential integrity was enforced between the ID column in Customer Info and the ID columns in the other tables.
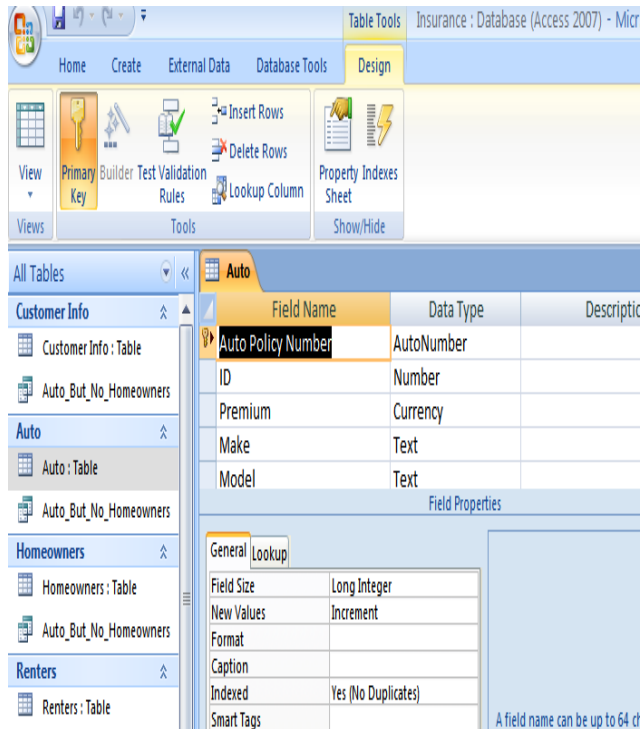

Figure 2. Column names for the Auto database table.

Appropriate data types for each table column were added in design view using the Data Type drop down arrow.

How might datamining be used to analyze this data? One way would be to identify customers who have Auto policies but not homeowners policies, and then target such customers with marketing efforts to sell them homeowners policies.

In SQL, the command would be:

SELECT [Customer Info].ID, [Customer Info].Last_Name

FROM ([Customer Info] INNER JOIN Auto ON [Customer Info].ID = Auto.ID) LEFT JOIN Homeowners ON [Customer Info].ID = Homeowners.ID

WHERE ((("ID") Is Not Null) AND ((Homeowners.ID) Is Null));

This query was created under Create | Query Design. The SQL was generated by selecting SQL View.
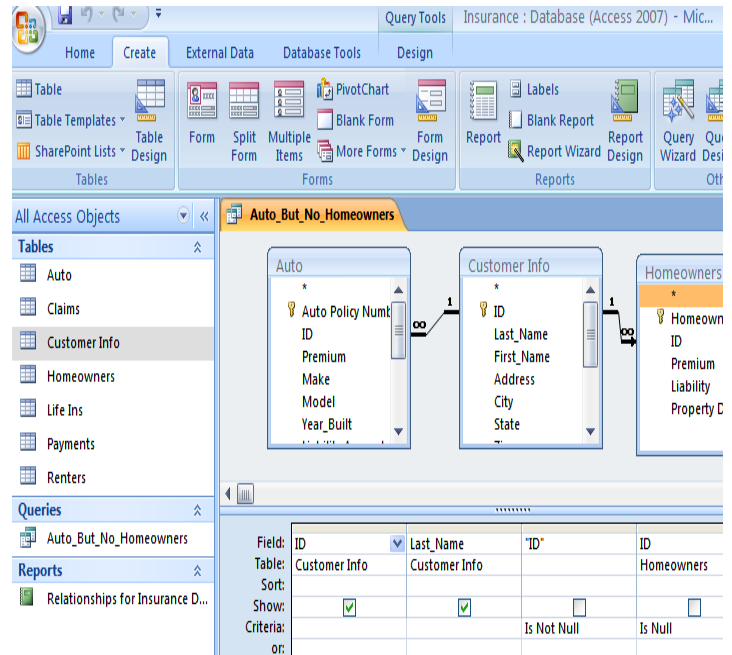

Figure 3. Building a query in Access.

Other uses of the database information would include finding the optimal setting for premium amount, amount of discount for multiple policy holders, and probability of a policy holder making a claim. Data miners might consider if likelihood of claim varies by region, climate, or driver's age or driving record.

There are several other database design exercises students might be asked to complete. Students could be asked to design database tables for an online retailer such as Amazon or for an online movie review website or an airline reservation website. Students could be asked to design a database for a credit card company, and include tables for customer credit history, bankruptcy history, credit applications, potential customers with their credit scores, formulas for the amount of credit to extend, formulas for which potential customers to contact with marketing literature, and transaction data on credit card balance payment records. Students could also be asked to design a database for a computer chip manufacturer, and include database tables for manufacturing data on yield ratios for various chip designs, the use of feed forward and feed backward process control, and total volume of chips produced compared with fab utilization rates. This data could be combined with sales data for various computer chips, cost data for supplies, and pricing data on various products.

.