Project: Data Design   **CS157A Fall 2020**   **Points: 50**

Problem statement. Your consulting company has been asked to develop a relational database design for a growing company that sells homeowners insurance. You gather the following information:

- Each customer of the company owns at least one home.
- Each home has associated incidents that are recorded by the insurance company.
- An insurance policy can cover one or multiple homes.
- The policy defines the payments associated with the policy, and a policy that covers multiple homes will show the payments associated with each home.
- Associated with each payment is a payment due date, the time period of coverage, and a date when the payment was made.

Process. Work in a team of up to 4 people. Perform and document each of the following steps:

1. Develop an ER diagram that captures the conceptual model. Show all detail in the ER diagram, such as the mapping and participation cardinalities. Identify any constraints not captured in the ER diagram. I suggest you follow the method for developing an ER diagram we covered in class.
2. Map the ER diagram to a relational database schema, using the procedure we have learned in class. Write down the database schema.
3. Normalize your relational schema using the procedure we learned in class. As a first step, document any functional dependencies that would indicate the need for normalization. Note that you are not required to normalize, but if your relation schemas are not in BCNF or 3NF then you should explain why you did not normalize. The final relational schema must be documented in SQL as create table statements.
4. Think of at least 5 questions about the data that would be interesting to company that sells Homeowners Insurance. Write each of them in SQL. Do not use only questions that are simple to express in SQL.

Report. After you have completed these steps, write up your work in a report. The report should clearly document the input and output of each step, the design issues you faced at each step, and how you resolved them. I would expect that your final report will be about 5-10 pages long (including figures).

I would expect the structure of your report to be something like this:

- Title, date, authors
- Introduction
    - one or two paragraphs explaining goals
- ER model
    - one or two intro paragraphs, ER model, parts of the conceptual model that couldn't be put in the ER diagram

- Relational schema derived from the ER model
  - one or two intro paragraphs, the DB schema, and paragraphs as needed for discussion
  - make sure to document constraints, including functional dependencies
- Normalized relational schema
  - one or two intro paragraphs, and then the final normalized relational schema in SQL
  - if you decide your relational schema from the previous step does not need normalization, then this section is not needed
- Sample data and SQL queries
  - some intro paragraphs explaining the questions that are of interest
  - SQL statements to express the queries ○ sample data written in SQL and the results of the queries on the sample data
- Conclusions
  - a couple of paragraphs that briefly summarize what you've done

SQL. You must also submit an SQL file, which contains

- SQL table create statements that define your database schema, including primary keys, foreign key constraints, any views or triggers you define, etc
- SQL insert statements for your sample data
- your sample SQL queries

I should be able to run your SQL file by pasting it into SQLite.

Grading. The project is worth 50 points.

Important notes.

- The work is to be done completely by your team -- please do not share anything with other teams or ask other teams to share with you.
- You work is to be submitted as a single PDF document. Only the team manager will submit the document.
- Do your work as if you are a professional consulting team.