Design outline:

1. We hope to transform the original denormalized dataset into a normalized database with multiple tables.
2. We try to encode categorical data as possible as we can for less storage and the ease of indexing.
3. We notice that there's hierarchical relations between geographical variables: city→ county→ state, also there's a column 'location' containing redundant information about city and state. Therefore, we dropped the 'location' column, and built the 3 tables with hierarchical relations.
4. Since there are lots of criteria columns with the same rating range, we extract those columns as a new table. Each criterion can have multiple hospital records, and vice versa. Therefore, we created a Rating_Records table to form the many-to-many relationship. In this way, we don't need to store data for unavailable rating records.