

Project II: Design and Implement MongoDB

Requirements and background

We gathered listing data from Airbnb in San Francisco. We filtered and left the most related fields. We created UML class diagrams, ERD Crow foot notation, relational schemas, and generated data to create a relational database. We also perform SQL queries to learn about the properties of Airbnb rentals.

Rules:

Field business rules:

- There are 500 listings and 400 hosts in the database.
- Listings only exist in the following 29 districts in San Francisco, including Sunset District, Union Square, Haight-Ashbury, Telegraph Hill, Marina, Castro, Golden Gate Park, Chinatown, Fisherman Wharf, Financial District, Mission District, North Beach, Presidio, Tenderloin, Japantown, Embarcadero, Noe Valley, Nob Hill, Soma, Pacific Heights, Bayview, Mission Bay, Stonestown, Hayes Valley, Bernal Heights, Alamo Square, Lower Pacific Heights, Stonestown, Russian Hill.
- There are only 5 types of Host response time, including within an hour, within a few hours, N/A, a few days or more, within a day.
- There are 15 types of room property.
- 10 kinds of room type.

Relationship business rules.

- One listing has one room.
- Every host must have zero or more rooms. Each host has one response time, each host has a district.
- Many hosts can be in the same district, many neighborhoods can be in the same district.
- One neighborhood can have zero or more rooms and can have many listings.
- Each room matches one piece of review and has the price from \$80 - \$300 per night. There is also description of the bathroom and number of bedroom and beds. Many rooms can be of the same property type, many rooms can have the same room type.

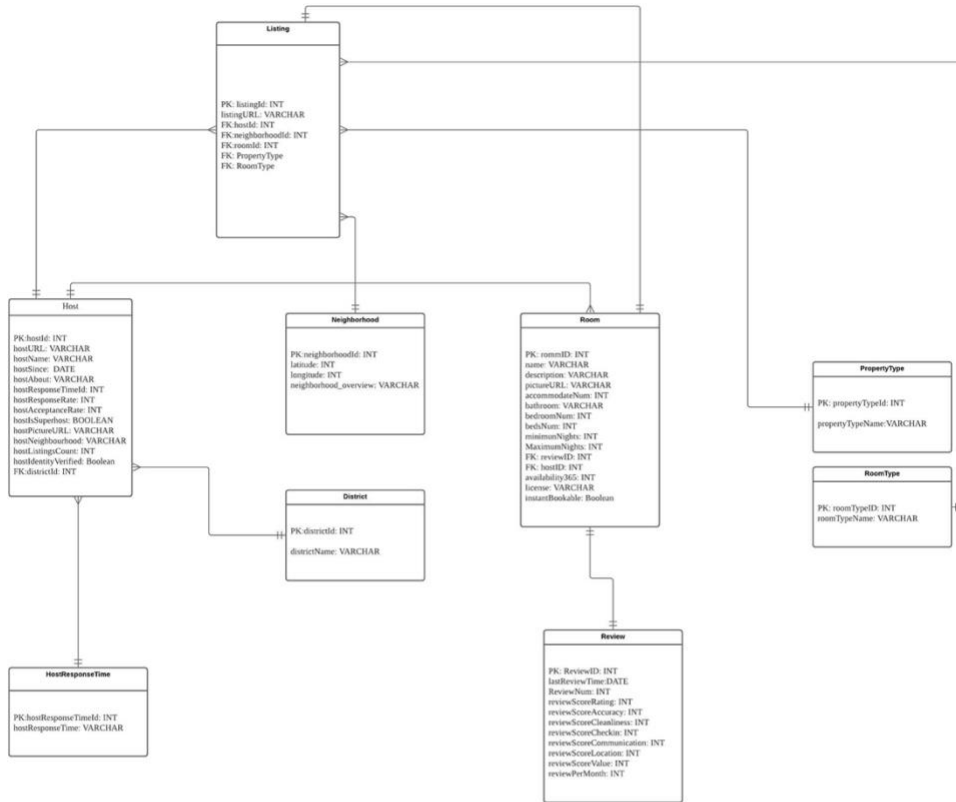
Nouns:

- Listing(Listings information include listing url, host, neighborhood, room, review)
- Host (including information like hostURL, hostName ,hostSince, hostAbout,hostResponseTimeId, hostResponseRate, hostAcceptanceRate, hostIsSuperhost,hostPictureURL, hostNeighbourhood, hostListingsCount, hostIdentityVerified, districtId)
- HostResponseTime(including hostResponseTime)
- Neighborhood(including longitude,latitude,overview,district)
- District(including names of different districts in San Francisco)
- Unit(describes relationship between listings and host)
- Room(including the room description, price, minimum/maximum staying nights, review about the room, and its propertyType and room Type)
- PropertyType(like private room in rental unit)
- RoomType (private room, entire apt/home, entire loft)
- Review(including the review about the room cleanliness, check in process, overall value, description accuracy, communication)

Actions:

- Has
- Belongs to
- Units

SQL



No-SQL

