The data is sourced from the Inside Airbnb website <http://insideairbnb.com/get-the-data.html> which hosts public available data from Airbnb. Particularly, we take the data of Boston area. The data comprises of three main csv files (all files are compiled on 13 Feb 2020):

1. Listings: detailed listing data showing 96 attributes for each of the listing
2. Reviews: detailed review data given by guests with 6 attributes. They are listing\_id, id, date, reviewer\_id, reviewer\_name and comments.
3. Calendar: details about booking for the next year by listing

----------------------------------------------------------------------------------------------------------------------------

For this final project, we are choosing to represent the data about Airbnb housing in Boston, including the information of rental space, host, customer and reservation. We are using SQL database to store our data.

To represent rental space, we collected the data of the name of each house, price for rental, number of bedrooms, number of bathrooms, description written by hosts, capacity of each property and average rating for this house which calculated by all stars given by reviews. We also kept track of the latest date of which host updated its information to let customers know how long it has been since the last update. In addition, we presented some pictures that demonstrates the surroundings and the house itself.

Also, we have five components which belong to certain housing: room type, property type, amenities, address and price. Among these, room type provides various types of room that customers can choose, such as entire home/apt, private room, shared room etc. Each room type can have zero or more corresponding rental spaces, while each rental space can only have one room type. Property type represents various house types, including apartment, guest suite, condominium, house, boat, townhouse, loft, bungalow, villa, bed & breakfast and others. Each property type can have zero or more corresponding rental spaces, while each rental space can only have one property type. Amenities indicate the services that supported by each house represented by Boolean values, such as wi-fi, parking, kitchen, heating and hot water etc. We have an amenity id to represent as identifier. Similarly, each amenity can have zero or more corresponding rental spaces, while each rental space can only have one amenity. Address contains the specific location of the house including zip-code, neighborhood, city, state and country where zip-code is the primary key. Each rental space is binding to only one address. Moreover, price involves price id, daily price, weekly price, monthly price, security deposit, cleaning fee, extra people where price id is the primary key of price.

Host is the owner of the house. Each host can host many rental spaces and has own identifier; however, each house must have only one host. We collected url, name, since, total listing and is verified of each house owner. What’s more, we cared pictures and whether a host is super host (friendly and nice) to offer a more customer-friendly service.

At the customer perspective, customers should have unique identifier as primary key as well as phone, email, date of birth, pictures to prove their identities. Customers can browse multiple houses that they desired. Additionally, customers make reservations to book rental spaces.

For each reservation, there is information about start date, end date, number of guests and booking date. Each reservation has its own id number and total price calculated by corresponding data.

Lastly, customers can write many reviews to give feedbacks and rate properties to these housing that they have lived in. Each review contains its own identifier. A review can only comment on one property; whereas each rental space can take advice from zero or more reviews.

----------------------------------------------------------------------------------------------------------------------------

We use SQL storage.

----------------------------------------------------------------------------------------------------------------------------

We use python as host language and SQL. There isn’t any machine restriction for our project.

We are using mysql.connector library as a bridge to connect python to SQL server, and some common libraries, such as numpy, matplotlib, etc,. We are using ktinter as frame to build our UI.

----------------------------------------------------------------------------------------------------------------------------

Airbnb, short for AirBed and Breakfast, has developed rapidly in recent years. It is a website providing new short-term housing which different from traditional hotel. This new type of housing provides people with more choices and experiences. Also, the price is a big advantage over hotels, so it is liked by a large number of travelers, especially young generations. We, as the users of Airbnb, want to discover more details of it in order to get a better user experience and introduce it to more people in need.

----------------------------------------------------------------------------------------------------------------------------

A close up of a map

Description automatically generated

A screenshot of a video game

Description automatically generated