Mid-term Report

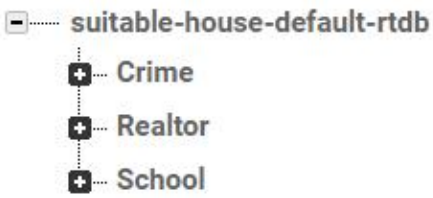
Title: Find the most suitable house

Milestone:

Collect data by API DONE
Collect some data from websites DONE
Project Proposal DONE
Data cleaning and processing DONE
Upload data to firebase DONE
Project mideterm report DONE
Build interface ONGOING
Connect with firebase ONGOING

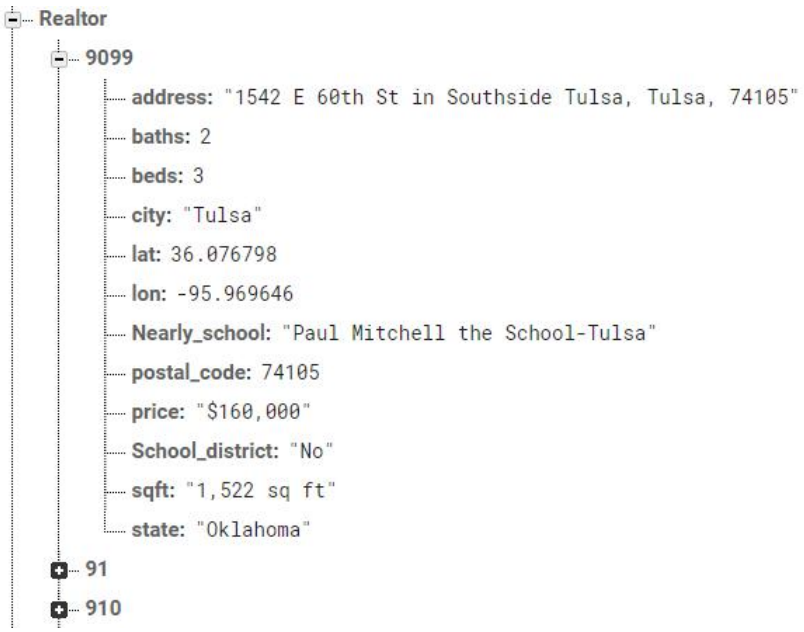
Dataset Collection

My project has 3 data sets. Some are from the API, others are from the website. I collect all data first, then I need to data cleaning and pre-processing. For example, I need to determine whether the school is a district house by calculating the distance between the school and the nearby school. When I done it, I save them as csv files. Then, I convert them to json file so that they can upload to firebase realtime database.
Here is the whole dataset structure:



1. Realtor dataset

From: real estate agent API
Dataset attributes: estate address, price, property type, area, neighborhood_name, advertiser_id, lead_forms, photo, facility, update time, bed , bath , lat, lon, sqft , state etc.
Dataset process: This dataset has too much information for me, such as facility, neighborhood_name, advertiser_id, lead_forms etc. First, I need to pick some useful attributes. Then, I just keep these attributes in the dataset. For some attributes, there are some outliers, and I also need to deal with them so that they are easy to use in subsequent steps.
Here is this dataset in the firebase:



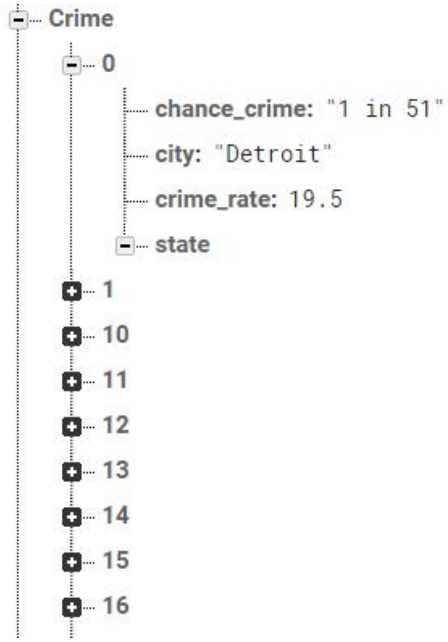
2. Crime rate dataset

From: scrape from website

Dataset attributes: chance_crime, city, crime_rate, state etc.

Dataset process: This dataset is fine for saving, because I can use all infomation. I just need to save it to the firebase database. I can use it to interface in the later step by matching the city with estate. For example, if this estate in the 'New York', I also can see the infomation about this city crime rate and chance crime in the interface.

Here is this dataset in the firebase:

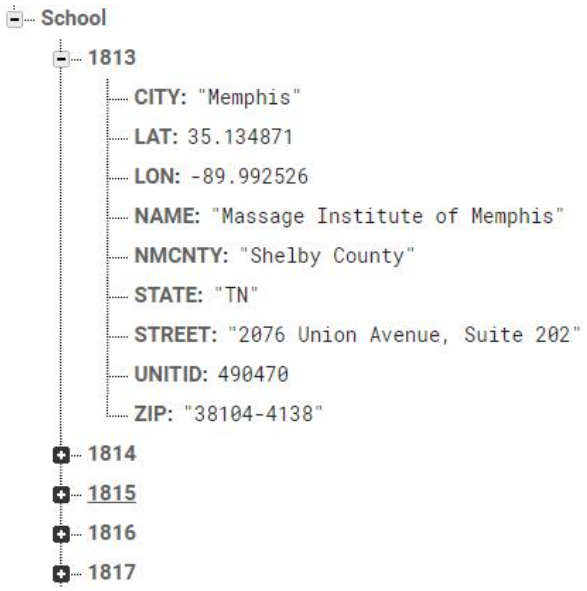


3. School infomation dataset

From: download from website

Dataset attributes: city, state, lat, lon, school name, street, unitid, zip etc.

Dataset process: Even the attributes of this data set are useful for my project, but some schools are useless in future processing. Therefore, I filter the data by all real estate cities to reduce the number of schools in this data set. After that, all schools can be used to calculate the distance to make a standard for judging whether it is a school district room. When I calculate all the distances of nearby schools from each real estate, I only choose the minimum distance, and then keep this school in the realtor dataset. I also set a range of 10 miles. If the distance is less than 10, it is a school district room; otherwise, it is not.
Here is this dataset in the firebase:



UI creation

For the UI creation, I want to create a html to display all the datasets. I just give a demo for later building. I want to achieve the goal that user can search the wanted real estate by some filters, like zipcode, city, price etc,. After that, users can see all statisfied real estate. Then, user will determine whether they want to buy it by other coditions. Like crime rate and whether it is a school district room.
Here is a demo for UI creation:

In order to achieve this goal, I need to connect firebase with html. I search some information online, they can give me some hint for finishing this part.

This article will get you up and running in five simple steps:

1. Create a project
2. Create a HTML file
3. Configure your Javascript connection
4. Connect
5. Open your web browser