

CS122 Project - Apartment Neighborhood Rating System

Mid-project Check-in #1 Memo

Team Han-Ji-Ji

Work proposed in timeline:

I. Data gathering and cleaning from data sources.

1. Chicago crime data ✓
2. Noise geographic information ✓
3. Neighborhood venue information ✓

II. Data storing and joining data from different sources. ✓

Work Delivered:

I. Data Preprocessing:

1. Dataset “crimes_one_year_prior_to_present” from city of Chicago data portal.
 - a. Removed duplicates and crimes with no location info.
2. Collected noise information by crawling howloud.com.
3. Used Google maps API to create helpful functions.
 - a. `get_latlng()`: function that returns the latitude and longitude tuple from a given address string.
 - i. The input address string is one of the user input.
 - b. `local_search(address, radius, venue_type)`: function that finds nearby neighborhood venues.
 - i. Originally we planned to use city of Chicago’s Park, Beach, grocery stores datasets, but the google API provides a better result. It allows us to look up a wider range of neighborhood venues in real time.
 - c. `query_crime_within()`: function that find crimes within a radius from the apartment address by querying our SQL database.

II. Data storing and joining data from different sources:

Created a SQL database, “geo.db”.

1. Table: crime
 - a. Schema (total of five): Case ID, date of occurrence, Primary Description (types of crime), latitude, longitude.
 - b. Since this is the only datas so far, we don’t have the need to join anything.