Real Estate Information Database System

*Li, Qiaoran*

*Yang, Zhaoxiong*

*Zhang, Liang*

Table of Contents

[INTRODUCTION 2](#_Toc508455944)

[DATA SETS 2](#_Toc508455945)

# INTRODUCTION

Our project assumes several potential factors that affect a family’s house choice.

We apply the knowledge we learn in class to prove decision supports taking into consideration of the following factors.

We gather information from several websites, using web crawler (written with Python) and other techniques. Then clean, and import to our database.

Our database includes several tables, such as house, crime rate, zip code and neighborhood relation, school, etc. Based on the database we created, we use SQL query to provide diversified information.

# DATA SETS

Most of the datasets are for Los Angeles

crime-rate: ﻿Zipcode, number of crimes

houses: HouseID, Zipcode, Address, City, State, Neighborhood, Price, Type, Beds, Baths, Build year, Space, Lot\_space, Price/sqft, Average\_Listing\_Price\_for\_zip, Median\_Sale\_Price\_for\_zip, Average\_price\_sqft\_for\_zip, description

population: Zipcode, Total\_population, Median\_Age, Total\_males, Total\_females, Total\_households, Average\_householdsize

private\_school: ID, Zipcode, Address, School\_type, Enrollment, FullTime\_teacher, etc.

public\_school: ID, Zipcode, Address, School\_type, Enrollment, FullTime\_teacher, etc.

zip-district: relations of zip code and district

zip\_lat\_long: relations of zip code and Latitude and longitude coordinates

# TECHNIQUES:

1. web crawler: we wrote web crawler using Python. This program is used to gather detailed house information from websites.
2. SQL database