Project group: 83

Team Members: Simran Singh Bapla, Forest Schwartz

Client:

http://flip1.engr.oregonstate.edu:2000/ – This is the URL for the main page
http://flip1.engr.oregonstate.edu:2000/admin – This is the URL for the administrator of the database with links in the side nav to the database information

Server:
http://flip1.engr.oregonstate.edu/9178 – Server is implemented on port 9178

Executive Summary:

During step 1 of the project we wrote an overview of the problem this database backend will solve. We also implemented our initial database outline and ERD. We received feedback from our peers to Improve database purpose description, enhance overview with more facts, restructure database for better alignment with final goal, and adjust 1:M relationships in outline. We then rewrote project overview and database purpose, restructured database: removed Sales, Customers, Rentals; added Cities, Zillow_estimates, Region_Statistics, and adjusted relationships in outline. During step 2, we normalized our database schema, and created a data definition .SQL file with sample data for all of our tables. We received feedback from our peers to remove crows' feet on relationships in schema diagrams, and commented code in DDL file. In step 3, we started implementing the UI for the project, and created a data manipulation .SQL file. We implemented most of the UI. We received feedback from our peers to add edit and insert functionality to pages, and a nullable relationship between two tables. We set up a nullable relationship between Homes and Zillow Estimates and added edit and insert buttons to the UI. In step 4 of the project we implemented CRUD functionality for Homes. We received feedback from our peers to change the homes table to display city names instead of the city id. In step 5 we implemented CRUD operations to the Cities table. We also received feedback to update the frontend to show changes after CRUD operations, which we implemented. During step 6 we implemented the rest of the CRUD operations for all the remaining tables, and wrote the final report.

Project Outline and Database Outline:

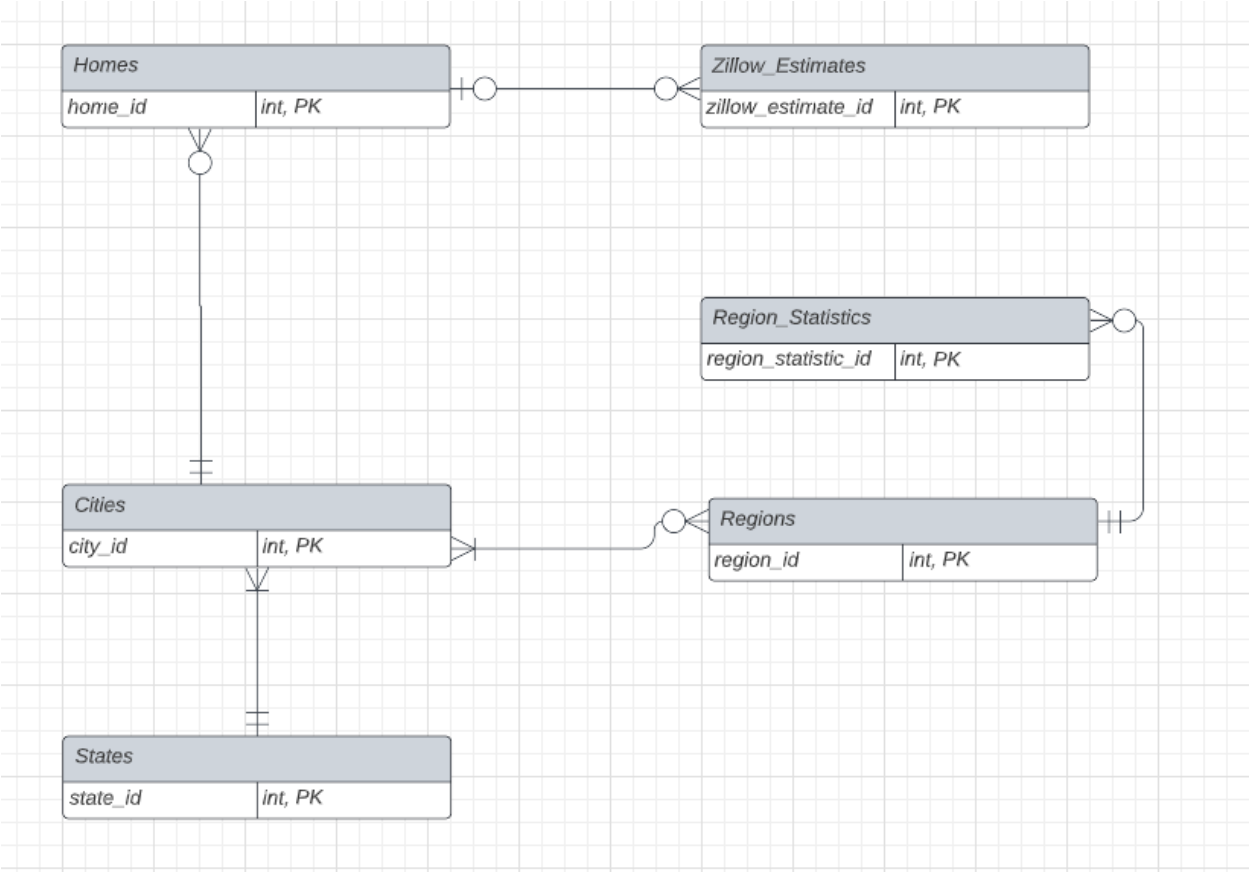
The goal of this project is to help real estate investors determine their next property to purchase. It does this by providing statistics on houses for each region obtained from Zillow. The database will store information about each house in America. Once a month, the Zestimate for each house will be recorded and stored in the database. The application will then calculate mean, median, and percent change statistics for each county based on the Zestimates obtained for each property at that specific date. The application can then use that information to show how the value of properties have changed over time

and with respect to different geographic locations. There are approximately 140 million houses in America. The long term goal would be to store information about each household available on Zillow.

- Homes: records the details of homes in the real estate investment application
  - home_id: int, auto_increment, unique, not NULL, PK
  - street: varchar, not NULL
  - zip: varchar, not NULL
  - sq_ft: int, not NULL
  - num_of_bed: int, not NULL
  - num_of_bath: int, not NULL
  - year_built: int, not NULL
  - lat: decimal, not NULL
  - lng: decimal, not NULL
  - city_id: int, not NULL, FK that references Cities
  - relationship: a 1:M relationship between Homes and Zillow_estimates, implemented with home_id as a foreign key inside of Zillow_estimates
  - Relationship: a M:1 relationship between Homes and Cities, implemented with city_id as a foreign key inside of Homes
- Zillow_Estimates: records of the zestimate (estimate) of a home scrapped from zillow.
  - zillow_price_id: int, auto_increment, unique, not NULL, PK
  - zestimate: decimal, not NULL
  - date: date, not NULL
  - home_id, int, FK that references Homes
  - relationship: a M:1 relationship between Zillow_Estimates and Homes is implemented with home_id as a foreign key inside of Zillow_Estimates
- States: Name of city
  - state_id: int, auto_increment, unique, not NULL, PK
  - name: VARCHAR, not NULL
  - Relationship: a 1:M relationship between Cities and States is implemented with state_id as a foreign key inside Citiest.
- Cities: Name of city
  - city_id: int, auto_increment, unique, not NULL, PK
  - city_name: VARCHAR, not NULL
  - state_id: VARCHAR, not NULL
  - Relationship: an M:M relationship between Cities and Region_Statistics is implemented with a separate table "Region_Statistics_has_Cities" with foreign keys named region_statistics_id and city_id inside of it.
  - Relationship: a 1:M relationship between Cities and States is implemented with state_id as a foreign key inside Cities
- Region_Statistics: Housing market statistics of a region of area
  - region_statistics_id: int, auto_increment, not NULL, PK
  - region_id: int, not NULL,

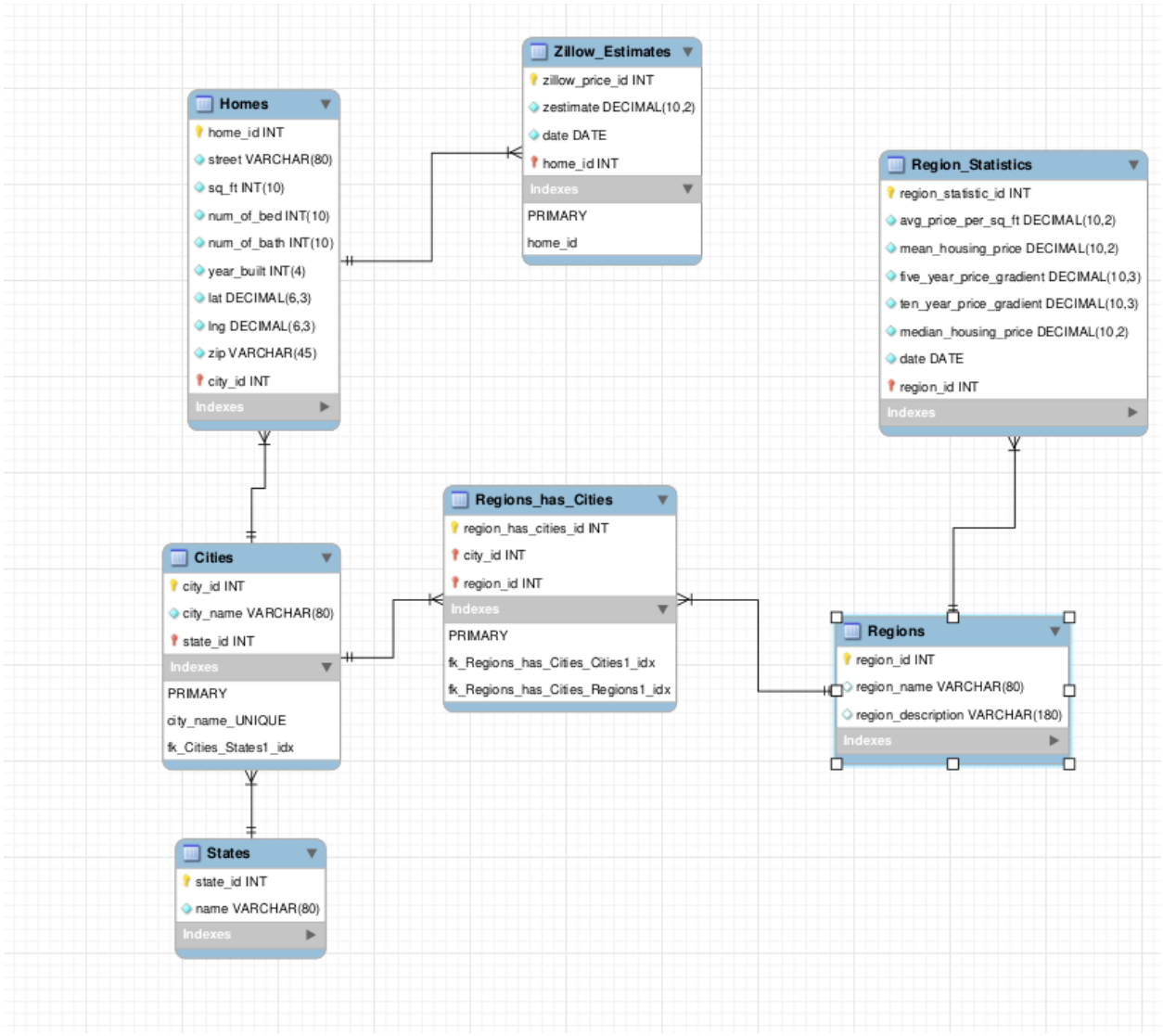
    - avg_price_per_sq_ft: decimal, not NULL
    - mean_housing_price: decimal, not NULL
    - five_year_price_gradient: decimal, not NULL
    - ten_year_price_gradient: decimal, not NULL
    - median_housing_price: decimal, not NULL
    - 1:M relationship between Regions and Region_Statistics with region_id as a foreign key inside of Region_Statistics
- Regions: Housing market statistics of a region of area
    - region_id: int, auto_increment, not NULL, PK
    - region_name: varchar(80), not NULL,
    - region_description: varchar(180), not NULL,
    - 1:M relationship between Regions and Region_Statistics with region_id as a foreign key inside of Region_Statistics
- Region_has_Cities: intersection table between Cities and Region_Statistics
    - region_has_cities_id: int, auto_increment, not NULL, PK
    - city_id: int, not NULL, FK that references Cities
    - region_id: int, not NULL, FK that references Regions
    - Relationship: M:1 relationship between Cities and Region_has_Cities with a foreign key city_id inside of Region_has_Cities
    - Relationship: M:1 relationship between Regions and Region_has_Cities with a foreign key region_id inside of Region_has_Cities

Entity-Relationship Diagram:



Schema:

1. First Normal Form (1NF): All the tables appear to be in 1NF as each column contains only atomic values, meaning each cell contains only one value.
2. Second Normal Form (2NF): Most of the tables appear to be in 2NF as each non-key attribute is dependent on the whole primary key, Third Normal Form (3NF): Some of the tables appear to be in 3NF as there are no transitive dependencies


Zillow_Estimates
zillow_price_id INT
zestimate DECIMAL(10,2)
date DATE
home_id INT
Indexes
PRIMARY
home_id
Homes
home_id INT
street VARCHAR(80)
sq_ft INT(10)
num_of_bed INT(10)
num_of_bath INT(10)
year_built INT(4)
lat DECIMAL(6,3)
lng DECIMAL(6,3)
zip VARCHAR(45)
city_id INT
Indexes
Region_Statistics
region_statistic_id INT
avg_price_per_sq_ft DECIMAL(10,2)
mean_housing_price DECIMAL(10,2)
five_year_price_gradient DECIMAL(10,3)
ten_year_price_gradient DECIMAL(10,3)
median_housing_price DECIMAL(10,2)
date DATE
region_id INT
Indexes
Regions_has_Cities
region_has_cities_id INT
city_id INT
region_id INT
Indexes
PRIMARY
fk_Regions_has_Cities_Cities1_idx
fk_Regions_has_Cities_Regions1_idx
Cities
city_id INT
city_name VARCHAR(80)
state_id INT
Indexes
PRIMARY
city_name_UNIQUE
fk_Cities_States1_idx
Regions
region_id INT
region_name VARCHAR(80)
region_description VARCHAR(180)
Indexes
States
state_id INT
name VARCHAR(80)
Indexes

Sample Data:

States

state_id	state_name
1	California
2	New York
3	Texas
4	Illinois

Cities

city_id	city_name	state
1	San Francisco	1
2	New York	2
3	Berkley	1
5	Los Angeles	1
6	Chicago	4

Regions

region_id	region_name	region_description
1	Bay Area	Placeholder
2	New York	Placeholder
3	Texas Triangle	Placeholder

Region_has_Cities

region_has_cities_id	region_id	city_id
1	1	1
2	1	3
3	2	2
4	3	4

Homes

home_id	street	sq_ft	num_of_bed	num_of_bath	year_built	lat	lng	zip	city_id
1	123 Main St	1200	2	1	1990	37.123	-122.456	94107	1
2	456 Elm St	1500	3	2	1980	40.789	-73.012	10001	2
3	789 Oak St	1700	4	3	1970	34.567	-118.91	90001	3
4	246 Birch St	1300	3	1	2000	41.876	-87.654	60601	4

Zillow_Estimates

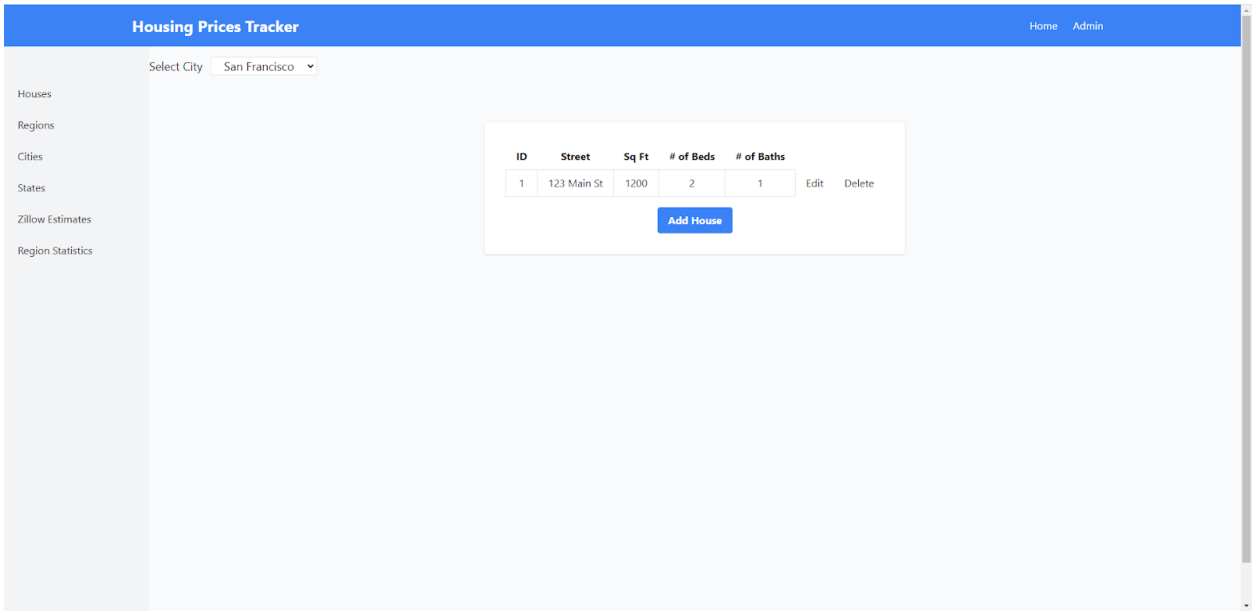
zillow_price_id	zestimate	date	home_id
1	2245681	2022-01-01	1
2	1245341	2022-01-01	2
3	1245481	2022-01-01	3
4	7245681	2022-01-01	4

Region_Statistics

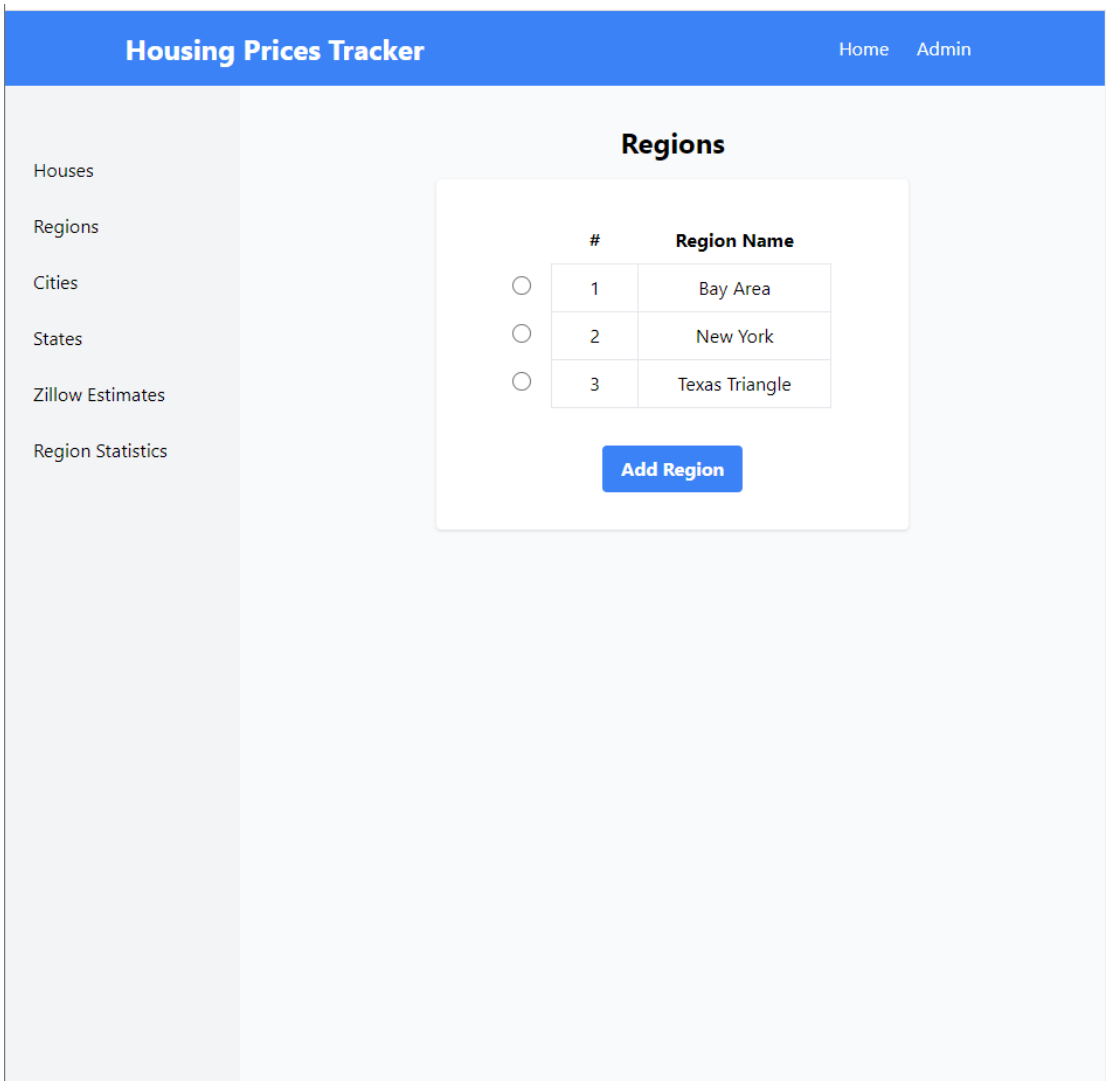
region_statistic_id	region_id	avg_price_per_sq_ft	mean_housing_price	five_year_price_gradient	ten_year_price_gradient	median_housing_price	date
1	1	500	800000	0.05	0.1	700000	2022-01-01
2	2	600	900000	0.06	0.12	800000	2022-01-01
3	1	400	700000	0.04	0.08	600000	2010-01-01
4	3	300	600000	0.03	0.06	500000	2022-01-01

Screenshots of Pages:

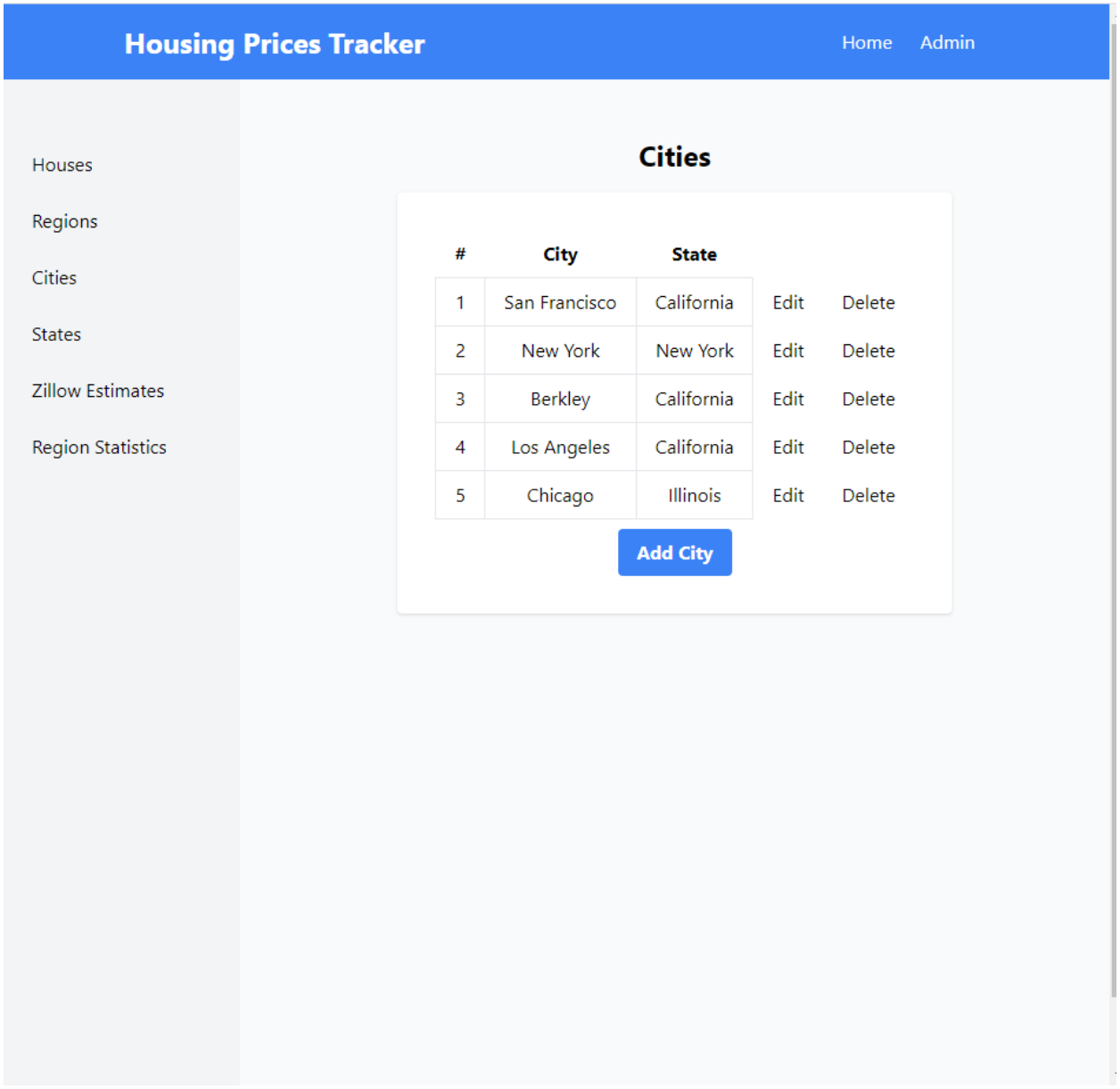
Houses (all CRUD steps implemented):



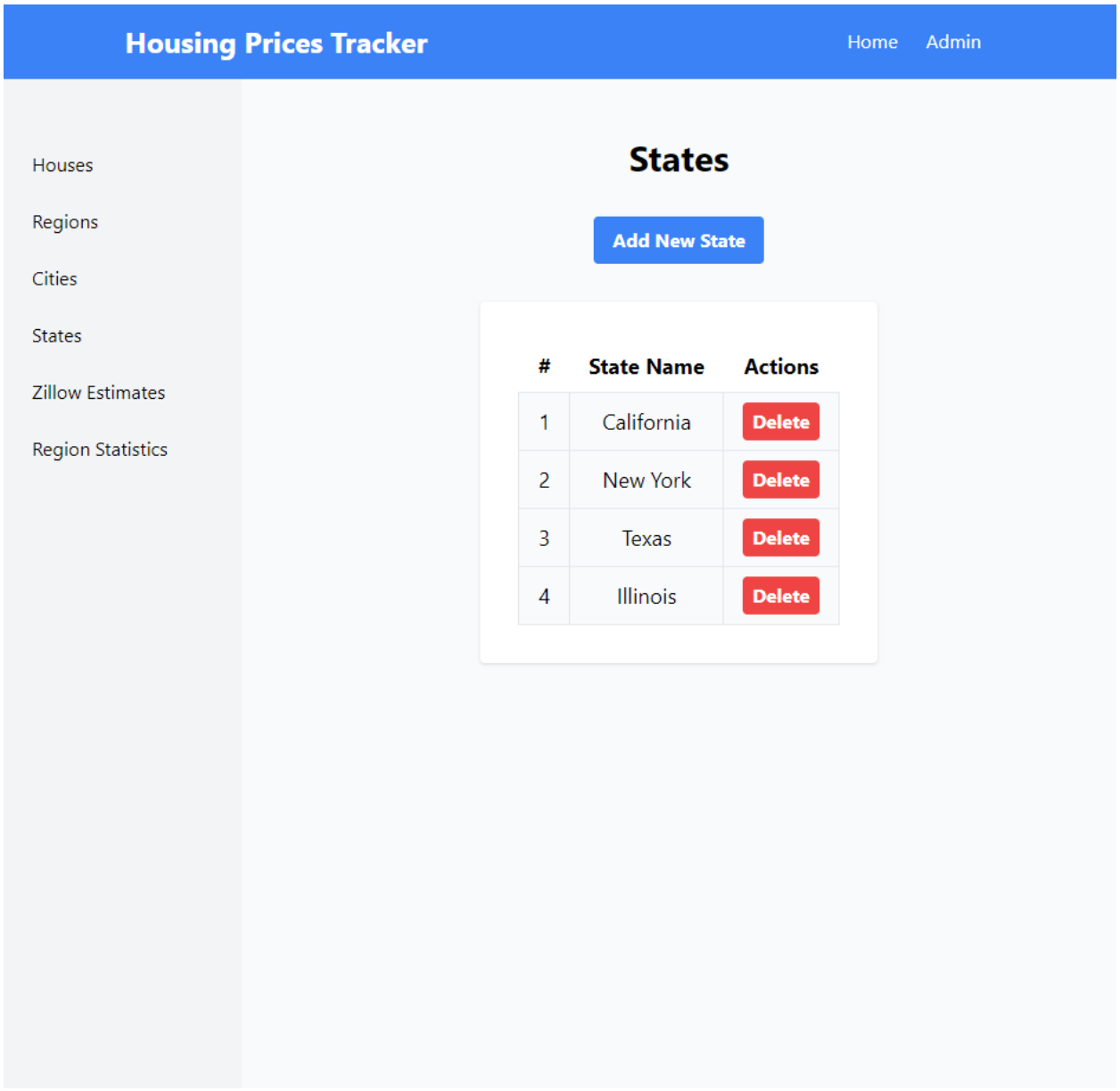
Regions (all CRUD functions implemented):



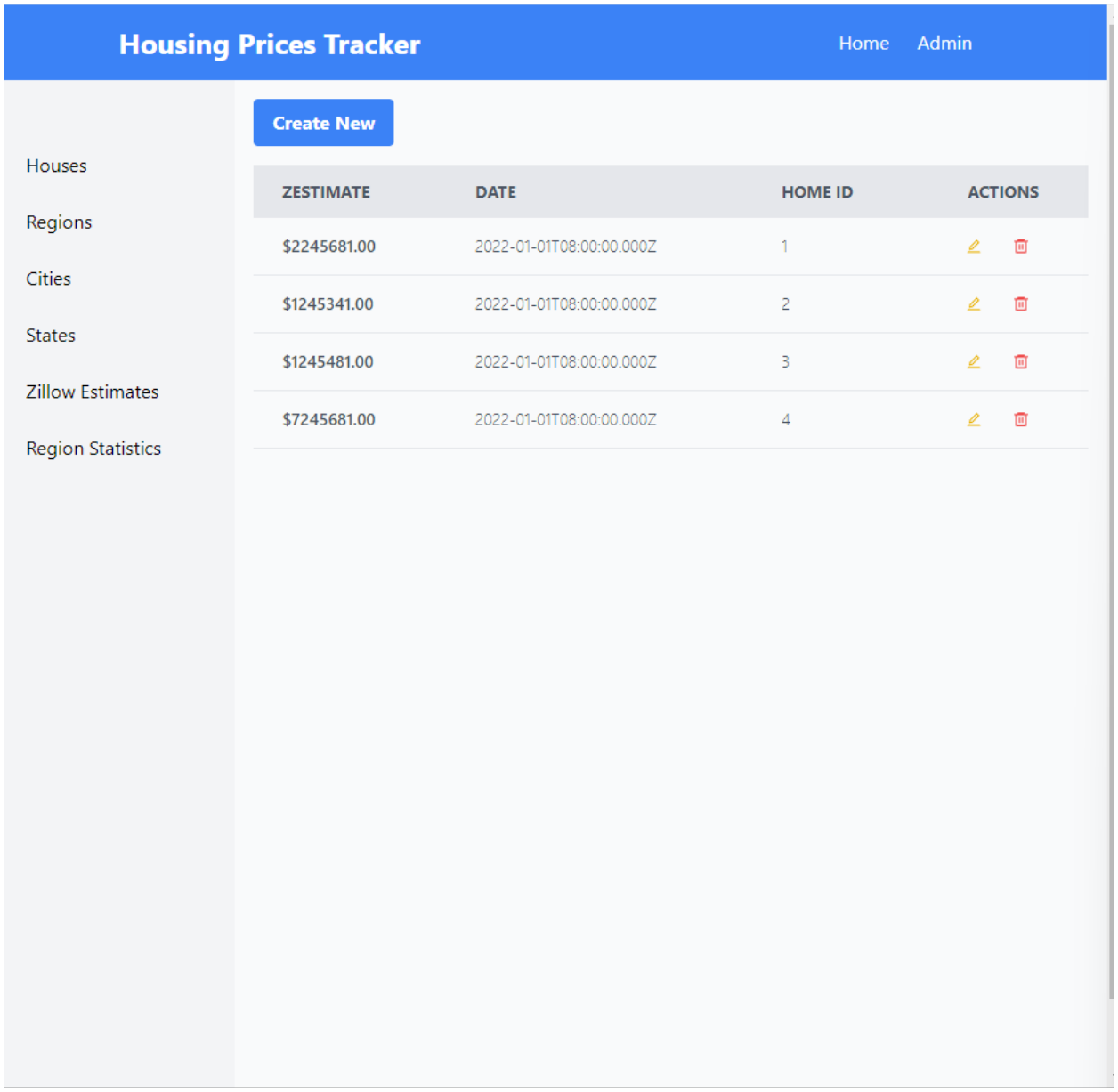
Cities (all CRUD functions implemented):



States (create and delete functions implemented):



Zillow Estimates (GET and DELETE Crud functions implemented):



Region Statistics (all CRUD functions implemented):