**Data**

Zillow and Yelp data were processed before inserting into the database (PostgreSQL). Data from other sources were scraped, cleaned and then added to the database.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Source | Dataset | Means of Fetching |
| 1 | https://www.zillow .com/research/data | Housing data | Downloaded csv files |
| 2 | https:// www.bestplaces.net/ | Crime rate, avg annual income, median annual income | Data scraping |
| 3 | https://www.schooldigger.com/ | Best schools and ratings | Data scraping |
| 4 | https://www.yelp.com/dataset/challenge | Restaurants and user ratings | Zipped tar file |

What are its characteristics (e.g., size on disk, # of records, temporal or not, etc.)?

Size on disk is 5GB

**Visualization**

Which UI we want to show in the poster?

Do we need to add Algorithm section?

**Approaches**

Most web applications provide users the ability to browse homes but don’t provide insight based on affordability and crime rate which the first two features aim at. The 3rd feature helps users view rising home price trends in states, counties, and zip-codes helping them make informed investment decisions. 3rd and 4th features haven’t been addressed by most of the researches in the real estate field. Most researches have been performed on home price pre-dictions without considering neighborhood amenities, which the 5th feature addresses but makes predictions on median home prices instead.

1. Find affordable states to live from a choice of all states in the U.S
2. Finds affordable and safer places to live within a state
3. Browse safer and affordable places in states, counties, neighborhoods within a specific median price range without going through the map view but using a graphical view
4. Make predictions on median home prices so users can invest in places that might have potential growth

**Problem**: It is always a daunting task to search for a new place to rent or buy. Budget and commute restrictions further make it cumbersome to find such places.

**Motivation**: We want to help users find desirable places faster and more efficiently using interactive visualization techniques, which can also help them see nearby amenities. We plan to highlight places where investments are worthwhile according to trends in the markets prices to help users make inform decisions

**Sanyam Khurana Nicholas Eagan Sai Padmaja Bhavani Guttikonda Justin Flick Viet Nguyen**

**Results**

References?

**Experiments/ Evaluation**

1. To experiment with D3 Graph algorithm for intuitively displaying a useful Graphical User Interface
2. To experiment and select an efficient algorithm to determine similar states and finding similarities between counties across states
3. To experiment and select the best ML algorithm based on prediction on accuracy for median home prices (regression), classification of safer place and affordable places
4. User testing to see if our UI is intuitive to the user or if any improvements need to be made

**Rental and Home Purchase Recommendation in the US**