Real Estate Market Analyzer

Basic Info

Shane Brown, ShaneJBrown@Gmail.com, u0852900
Sigmund Chow, sigmundccs@gmail.com, u0597938
My Huynh, my.huynh@utah.edu, u0729654

Project GIT repository:
https://github.com/shanejb/dataviscourse-pr-RealEstateMarketAnalyzer

Photo:
https://www.linkedin.com/pulse/beginners-tips-rent-commercial-space-office-space-bangalore

1. Background and Motivation

Renting has long since been an acceptable way to pay for housing throughout the civilized world for many, many generations, but does renting make financial sense? Part of the American dream is to own your own home, but is owning always the wisest financial decision?

In a recent article in Bloomberg it states that "Three in five retirees surveyed by the Transamerica Center for Retirement Studies said making money or earning benefits was at least one reason they had retired later than they planned to. Almost half said financial problems were their main reason for working past 65."[1] Millions of Americans are working longer and retiring later in life. The traditional wisdom is that if you work hard pay into social security and invest into 401K, or similar investments, that those will be sufficient for retire. Unfortunately for millions of Americans they are realizing that they do not have sufficient funds to retire. For some real estate investing is a great way to build wealth and provide additional funds for retirement.

2. Project Objectives

This project aims to answer the following questions. When does it make sense to buy? When does it make sense to rent? What areas in the United States are the best rental markets? What areas in the United States are the best real estate markets with the highest appreciation forecast?

The goals for this project are threefold. First, we would like to educate ours and others on what factors to consider before buying a property versus renting. Second, we would like to be able to display areas which are the best rental markets in the United States of America. Lastly, we would like to show which areas have the highest appreciation forecast.
The benefits to this project are to educate individuals on when it makes the most financial sense to rent and when it makes the most sense to buy. This will empower the individual to take control over their financial future.

Another benefit is to educate others on where the best rental markets are. This will be done from two different perspectives the renter and property owner. For the renter we will represent which areas has the best rental values. For the property owner we will represent which areas provide the highest rental income to property cost ratios.

---

[1] Steverman, Ben. "'I'll Never Retire': Americans Break Record for Working Past 65", www.Bloomberg.com, 13 May, 2016.
http://www.bloomberg.com/news/articles/2016-05-13/-i-ll-never-retire-americans-break-record-for-working-past-65

Lastly, by providing visualization tools displaying which areas have the highest appreciation forecast this can benefit both the property investor and individuals looking to buy a single home. This is because the higher appreciation the higher the ultimate return of their investment will be.

3. Data

Zillow has come up with their own methodology for real estate data - Zillow Home Value Index. Each ZHVI is a time series tracking the monthly median home value in a particular geographical region. The ZHVI are available for seven geographical levels: neighborhood, ZIP code, city, congressional district, metropolitan area, county, state and the nation. We will be using their ZHVI, Zillow Rent Index(ZRI), affordability data and some other metrics for this project.

Limitation:
Data for certain regions/cities are sometimes incomplete(e.g. Salt Lake is not in the median sale price list) and we understand that it's hard task to gather data on such a large scale. Zillow has also mentioned on that webpage that they are in the process of transitioning some data sources and expect the transition to be complete before the end of 2016.

Data Source Links:
ZHVI - http://www.zillow.com/research/data/#median-home-value
ZRI - http://www.zillow.com/research/data/#rental-data
Map boundaries data - https://www.census.gov/cgi-bin/geo/shapefiles2010/main

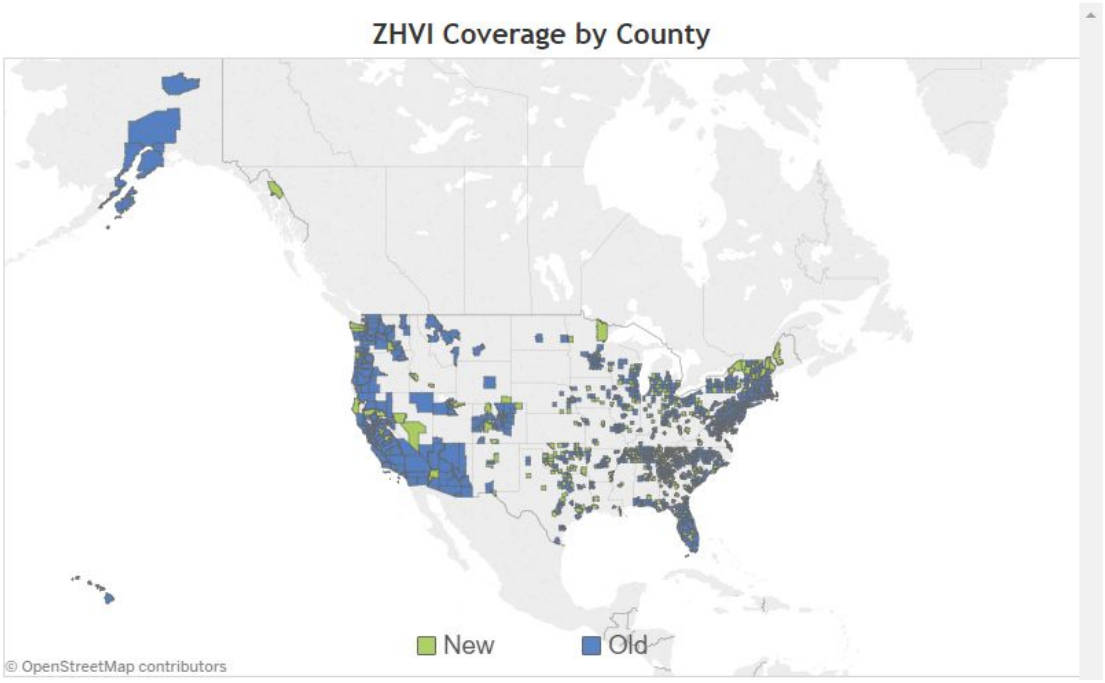


Image illustrates the new coverage of ZHVI (green) in addition to old coverage (blue).

4. Data Processing

Most of the data provided by Zillow actually come in CSV format. However, we will use Python to pre-process the data when necessary, mostly for aggregating, separation and reformatting the data to reduce real-time processing on the d3 side. For instance, the mortgage affordability and rental affordability data need to be separated from the same file.

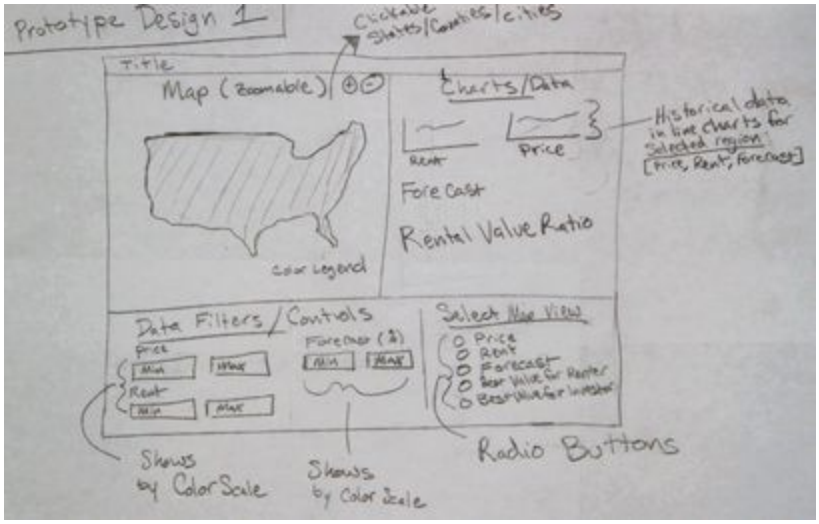
In order to implement US map, we need GeoJSON format suitable for working with D3. There are several things that could involve in processing:

- To convert shapefiles to GeoJSON, we need use ogr2ogr, part of the GDAL package.
- To convert TopoJSON to GeoJSON, we will use one of the method we previously used in one of the homeworks.

5. Visualization Design

A. Design 1

a. Sketch

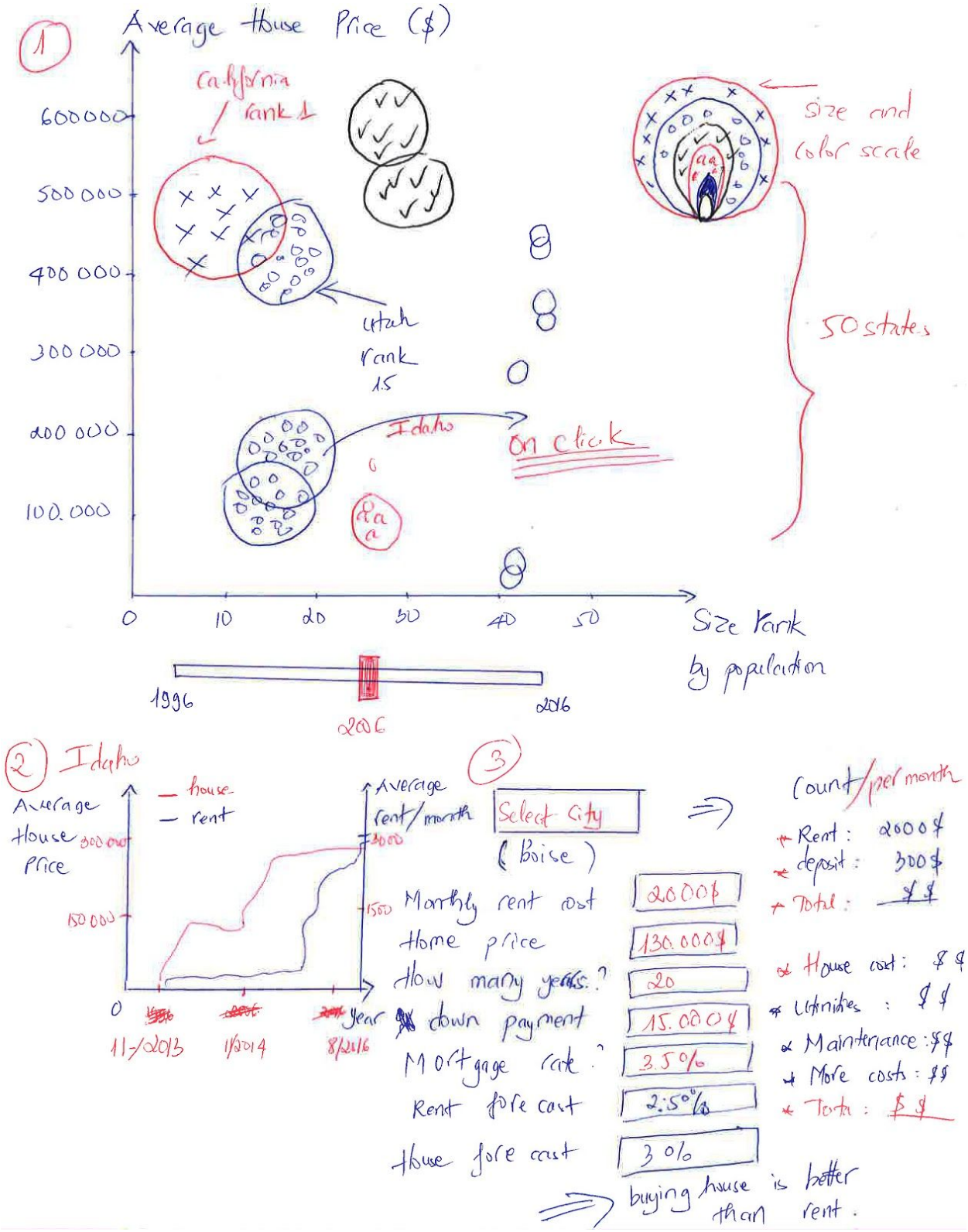


b. Description

Prototype design 1 has one portion of the display area to have a zoomable and selectable map of the United States. The user is able to single-click on an area of the map to select it and double-click to zoom in. The map will show a color scale of one of the selectable criteria, such as, price, rent, appreciation forecast, and best rental value. Another section will be to display charts and actual data values based on the selected region and selected filters. Lastly, a data filters and controls section will provide the user the ability to apply filters and controls such as to set ranges for different criteria and to set which view to display on the map. The color scaled map will give the user the ability to get a sense of what the market is like in one area compared to another. The charts and data section will provide the user with more solid numbers and visualizations which can represent actual values and historical data for the selected areas.

B. <u>Design 2</u>

a. Sketch



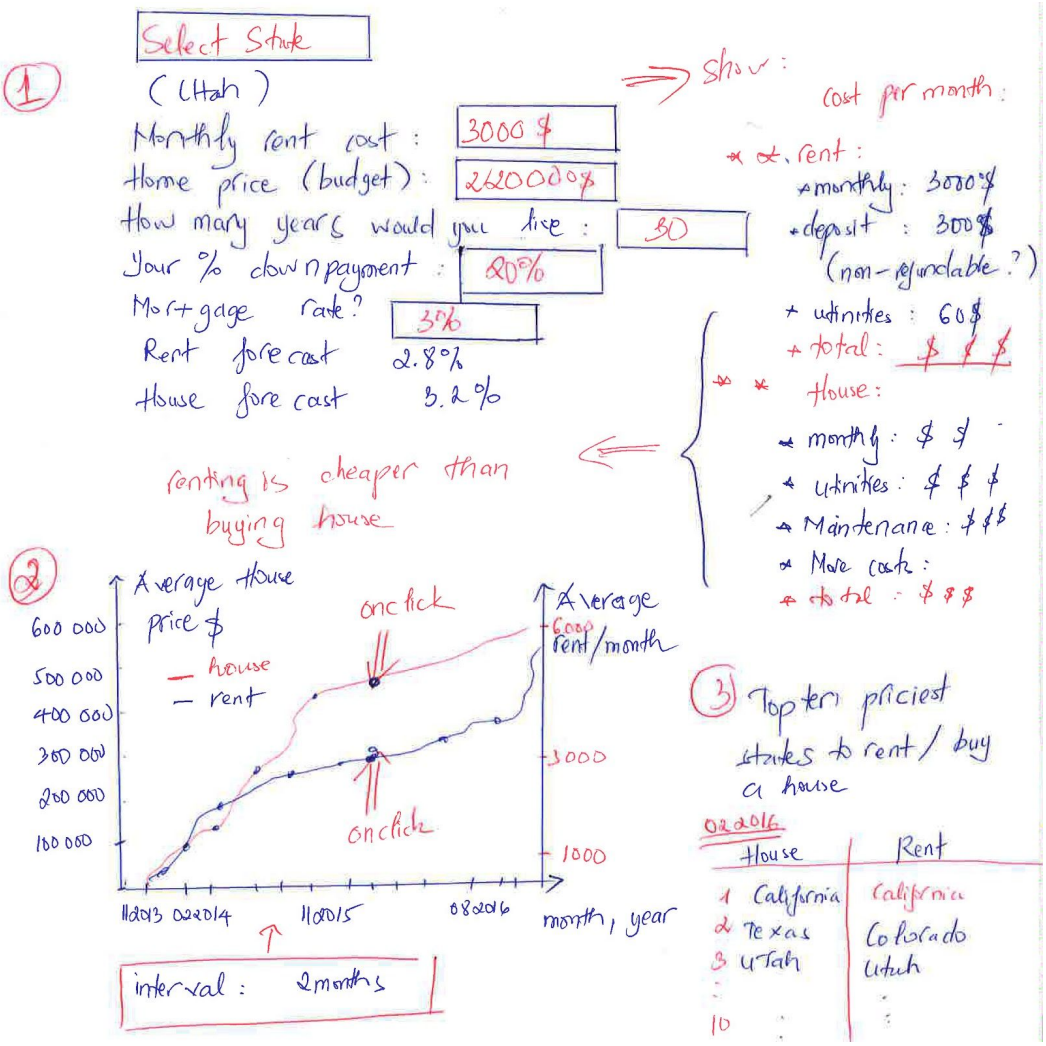
b. Description

There are 3 visualizations in this design. Details as follows:

- ***Datasets are used in order to implement the design include data of***
  - median house prices of all states in US from 1996 to 2016.
  - median house and rent costs for all states in US from November 2013 to August 2016
  - % of house and rent value forecast
- ***The bubble chart (1)*** shows how fast housing price increases over time. As we can see in this chart:
  - Time slider with a scale of 20 years from 1996 to 2016 in 1-year increment which allows users interact with bubble chart.
  - One bubble represents for one state. Size and color of each circle shows US state size rank by population. The more state's population is, the more size of circle is.
  - Above figure is an example when the button moves to 2006 what shows California has number one in state rank and highest average housing price.
- ***Line chart***: ***(2)*** by clicking on a bubble associated with the state, line chart ***(2)*** will display the data of rent and housing cost of that state. In above sketch, line chart is an example when users click on Idaho.
- ***Calculator (3)*** has a task to answer the question if buying house is always better than renting. It allows users select a city where they live or plan to move in. Users can enter their effort budgets and sensible options for renting and buying cost in order to see and compare how much they would pay per month for a house. In addition, line chart ***(2)*** is able to show rent and house cost raising over 20 years for associated city that users select in the calculate tool.

C. Design 3

1. Sketch



2. Description

- ***Third design also has 3 visualizations using datasets of***
  - median house prices of all states in US from 1996 to 2016.
  - median house and rent costs for all states in US from November 2013 to August 2016
  - % of house and rent value forecast

- ***Calculator (1)*** is assigned to answer the main question “renting vs buying a house: which is better?”. As in design 2, it allows users select a city where they live or plan to move in. Users can enter their effort budgets and sensible options for renting and buying in order to see and compare how much they would pay per month for cost of living. Calculator in design 3 is an example renting is cheaper than buying a house in Utah. Obviously, this is just a example which explains how we implement the calculator, not correct information.

- ***Line chart (2)*** provides data of average renting and housing price increases from November 2013 to August 2016 in 2-month interval. We select the period time 2013 to 2016 because we would like to show how fast the costs raise every 2 months. Otherwise, we have data available for all states since Nov. 2013.
- ***Table (3):*** dot makers on line chart ***(2)*** respond to click event in order to update table of top ten priciest states.

D. Final design

After we've finished three designs, we found some features that we should have in the final sketch. In other words, they incorporate the best of those designs. Our final sketch as follows:

1. **Sketch**

2. **Description**

   Slider View: This will be a time slider, as in design 2, for the years 1996 to 2016, in one year increments. This will be interactive by the user to update the rest of the views.

   Map View: The map section, as in design 1, will display the entire United States of America with the states outlines and names shown. The map will be interactive where the user is able to click on and select one or more states. Selected states data will be shown in the line charts view and detail view. The map will also have a color scale which will represent the rent statistics for the year selected by the slider view.

   Bubble Chart View: This will be as in design 2. Chart view has a function that shows how fast housing price increases over time. As the map, bubble chart will be interactive with time slider in order to show average house price of all states have changed in 20 years. One circle represents for one state. Size and color of each circle shows US state size rank by population. The more state's population is, the more size of circle is.

   Line Charts View: This will represent historical rent and home price data for each of the selected states in either the map view or bubble chart view.

   Detail View: This will represent more specific details of each of the selected states. Some details it may include, but not limited to, are population, average home price, average rent rate and so forth.

6. Must-Have Features

1. A Map that displays all states with boundaries

2. Rent throughout the years
3. Median home price throughout the years
4. Bubble chart shows appropriate data
5. Bubble chart and map interact with time slider
6. Detailed view of the selected states

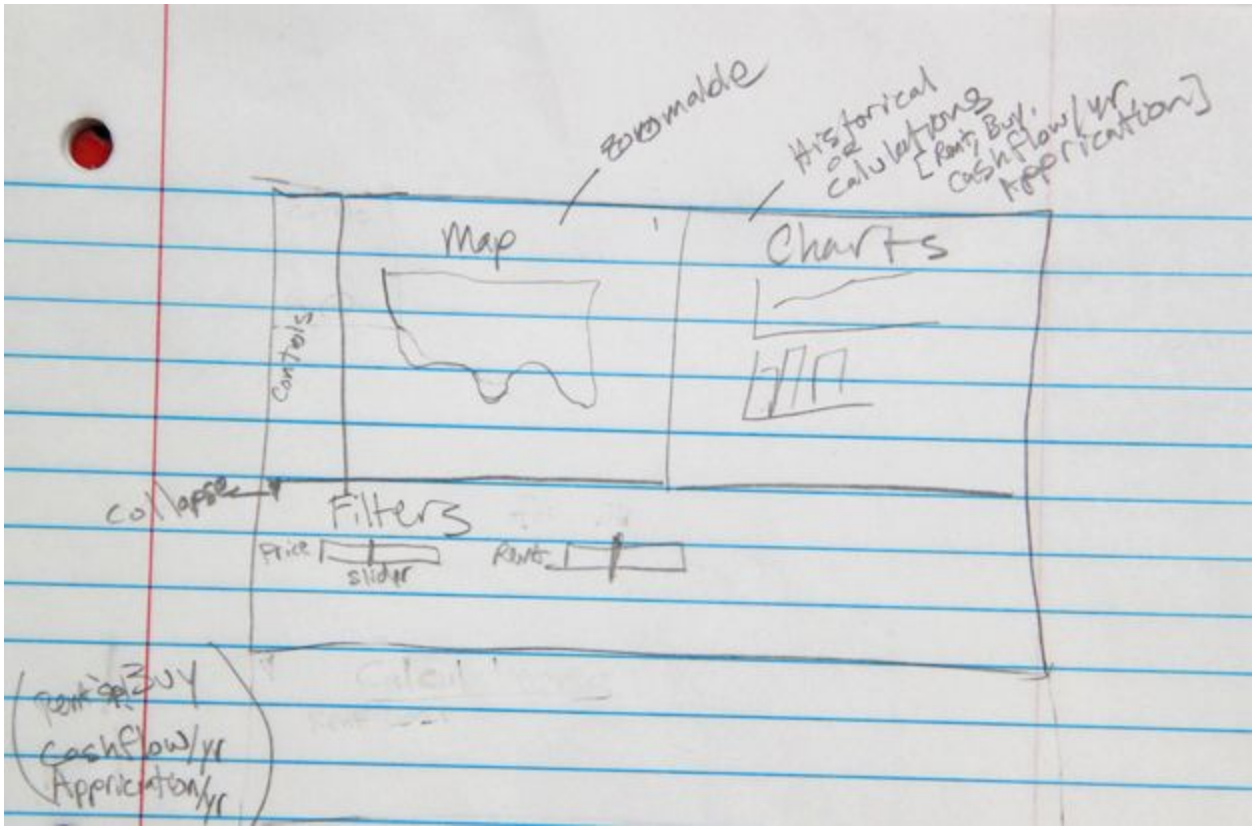
7. Optional Features

- Include visualizations for crime rate per capita of locations
- Include visualizations for school ratings of locations
- Include data on the cost of living of locations
- Include interest rates since 2011
- Add pan and zoom to the map

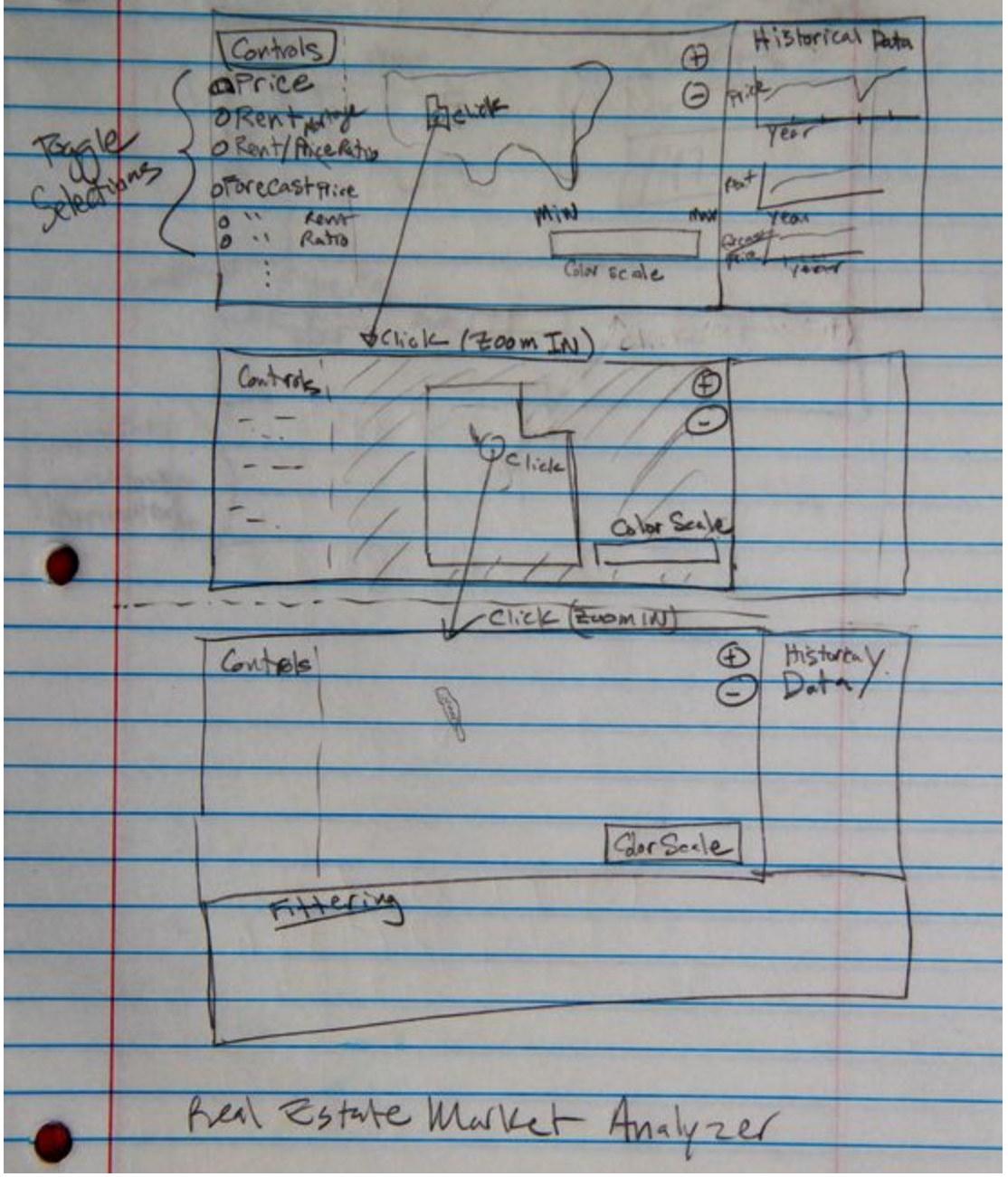
8. Project Schedule

Week	Work
Oct. 26th - Nov. 1st	Modify/add/remove features of final design after getting project peer feedback from an another group if necessary Complete outline design Complete datasets processing Complete classes, elements, and initial structure
Nov. 2nd - Nov. 11th (Milestone-1)	Complete data acquisition Must have all basic structures Visualization can shows the map, charts, views No interaction needed
Nov. 12th - Nov. 28th	Complete must-have features Connect views, implement interaction between map, bubble chart with slider Doing optional features if have time
Nov. 29th - Dec. 1st	Complete process book, screen-cast, final website
Dec. 2nd	Review all works Due date

Additional Sketches

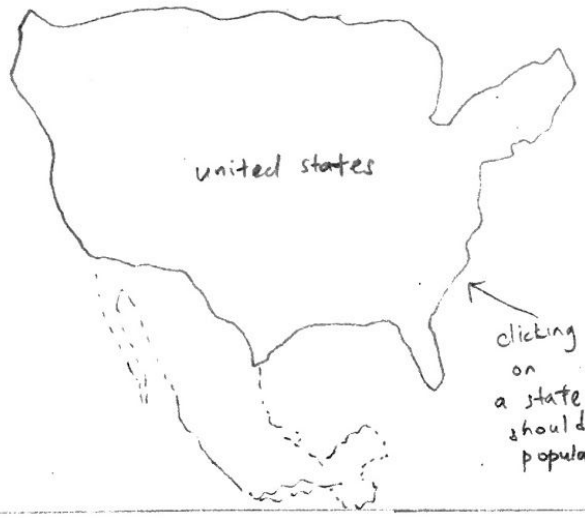


Objectives see which markets are best to invest in



Real Estate Market Analyzer

Map



Information / Data

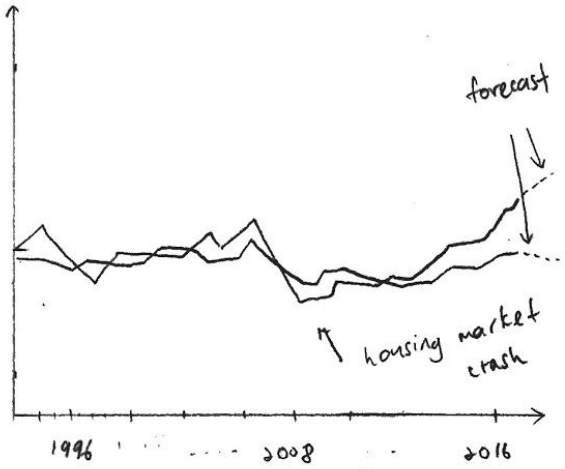
population

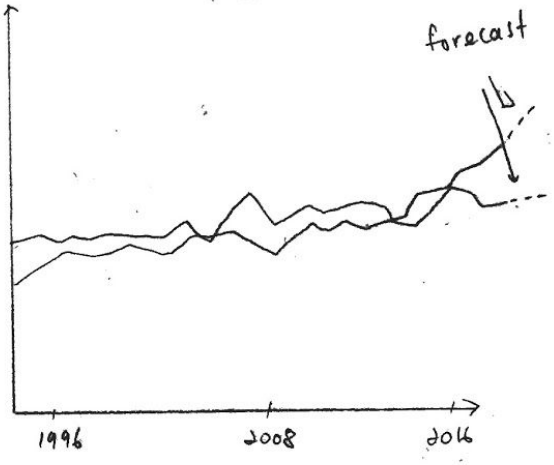
state median
highest
lowest

compare to [california ▼]

can select more than 1

Trends





Ideas:

1) ability to compare up to 5 states? that will show 5 lines
2) ability to compare with prediction?
3) compare ZHVI to sale price & list price?