Process Book for Vis Project

Yichu Zhou(UID:u1076093:, Email:flyaway1217@gmail.com)
Han Han(UID:u0615379, Email:h.han@utah.edu)

Background and Motivation:

The housing expense in the US has been increasing rapidly during the last five years, and significantly affects the living cost, especially in the Bay area and the Seattle metropolitan area where there are a lot of tech companies and have the largest job market for computer science graduates. Local housing prices has become one of the major concerns for the career choice of the fresh graduates who are going to enter the job market or want to start their own start-ups. As graduate student in computer sciences, we are interested in looking at the history of housing prices in different areas and what might be the contributing factors for the housing prices' changes.
Project Objectives:
1: What is housing prices for different locations in U.S. ?
2: How about the current housing prices compared with that in the history?
3: Which is the better choice now: owning or renting? Is it a good time for real estate investment?
4: What other factors contributing the housing prices?

Data:

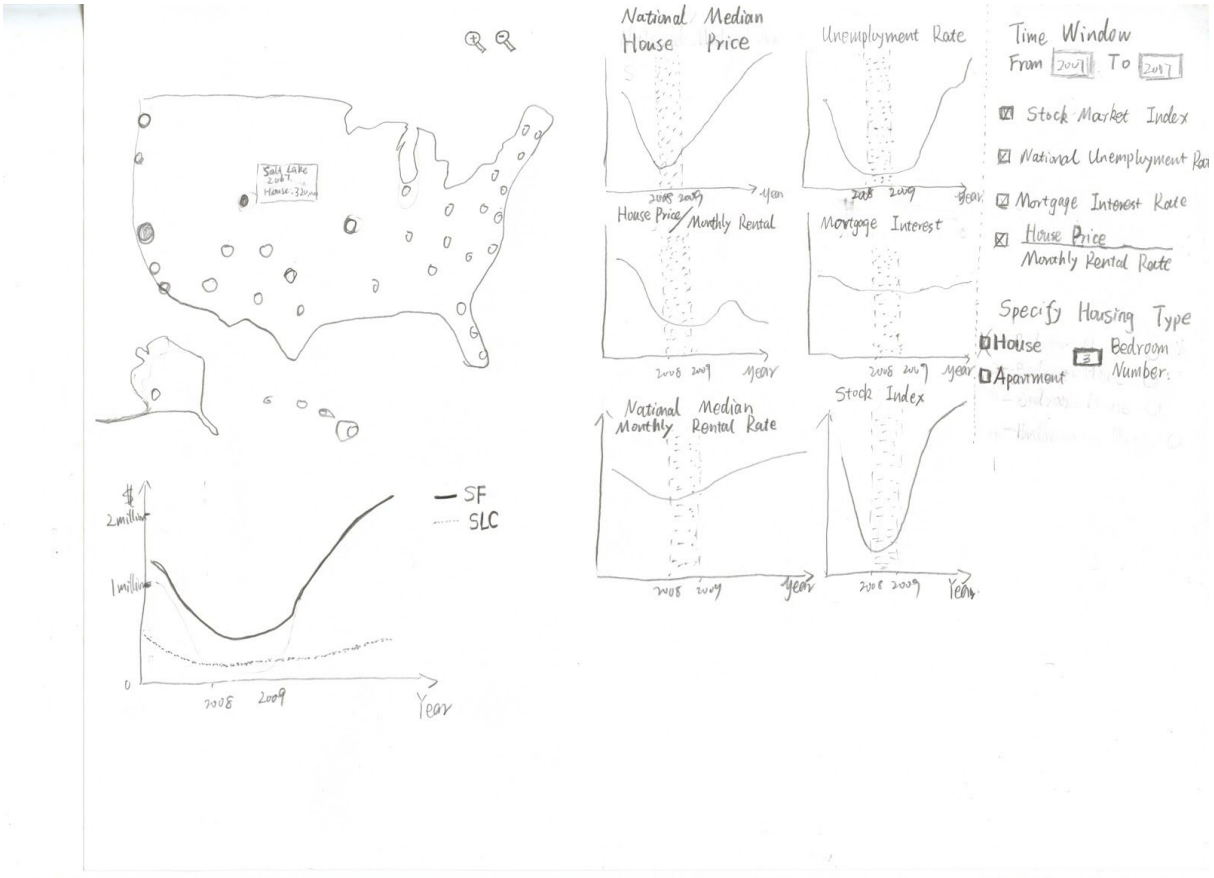
The housing price data will be from Zillow. (https://www.zillow.com/research/data/)
The stock market historical data will be from Yahoo Finance. The Dow Jones Industrial Average Index will be used.
The historical data of the unemployment rate in the US will be from the United States Department of Labor: (https://data.bls.gov/timeseries/LNS14000000).
The historical data of the mortgage rate in the U.S. will be from Freddiemac. (http://www.freddiemac.com/pmms/pmms30.html)

Data Processing:

Most of the data needed for the visualization will come from the downloaded data directly. To analyze the potential profit to own a house, one derived data will be used, which is the ratio of the House price divided by monthly rental rate.

Visualization Design:

Design 1:



There will be two major parts of this visualization: the left is the major area for the data visualization, the right will be the control panel.

On the control panel, there are several buttons and input windows, including:
1: Input window for input of the chosen time period. If no time period is specified, the default time window will the the last ten years(2007-2016).
2: Options for the visualization of: Stock market index, national unemployment rate, the mortgage interest rate, and the house price/monthly rental rate ratio. Without checking these options, the graph for these individual features will not be displayed.
3: Options to specify Housing Type, including House, Apartment, and Bedroom Number. Without specify the housing type, the default will be the median price for all houses/apartments.

In the major display area:
On the right, there will be graphs visualizing the National Houses Median Price and National Median Monthly Dental Rate during the input time window. If other options were chosen,

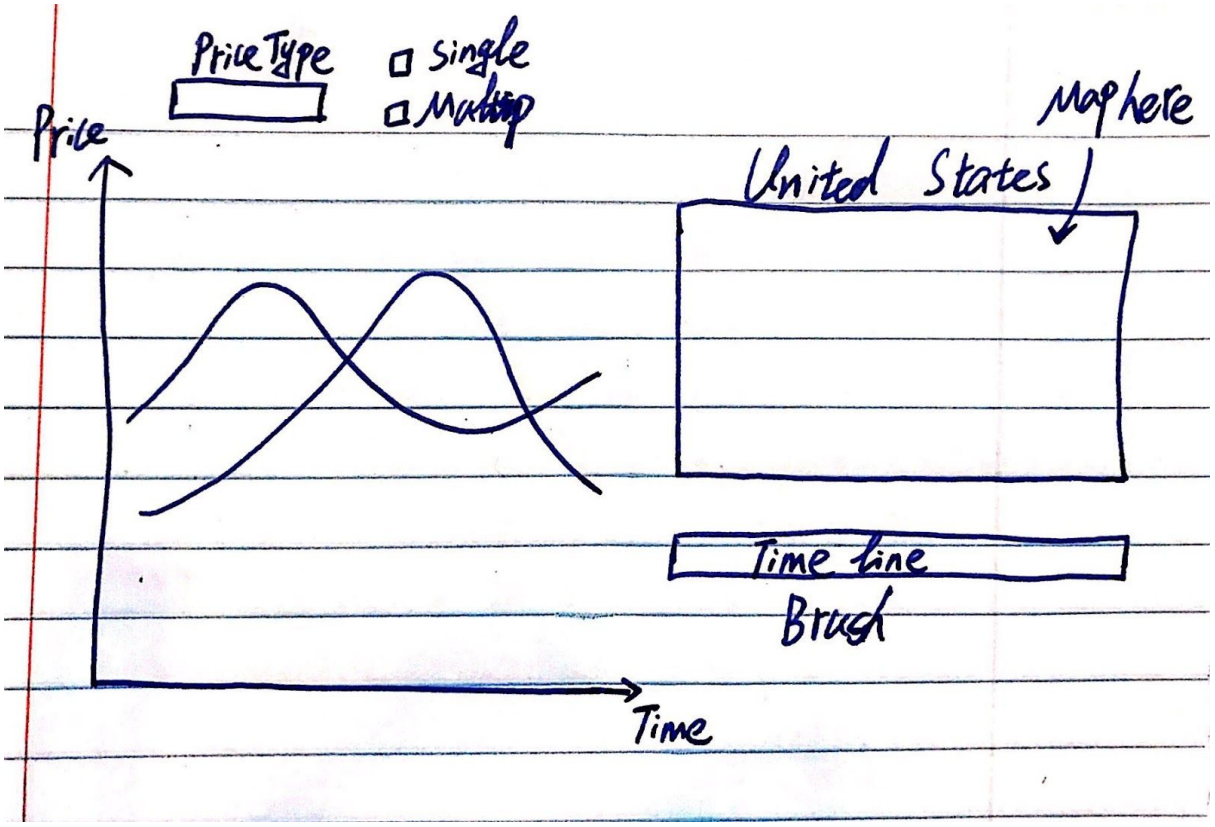
such as the unemployment rate and mortgage interest, they will also be visualized in this area. The mark for these data visualization will be lines, and the channel will be position.

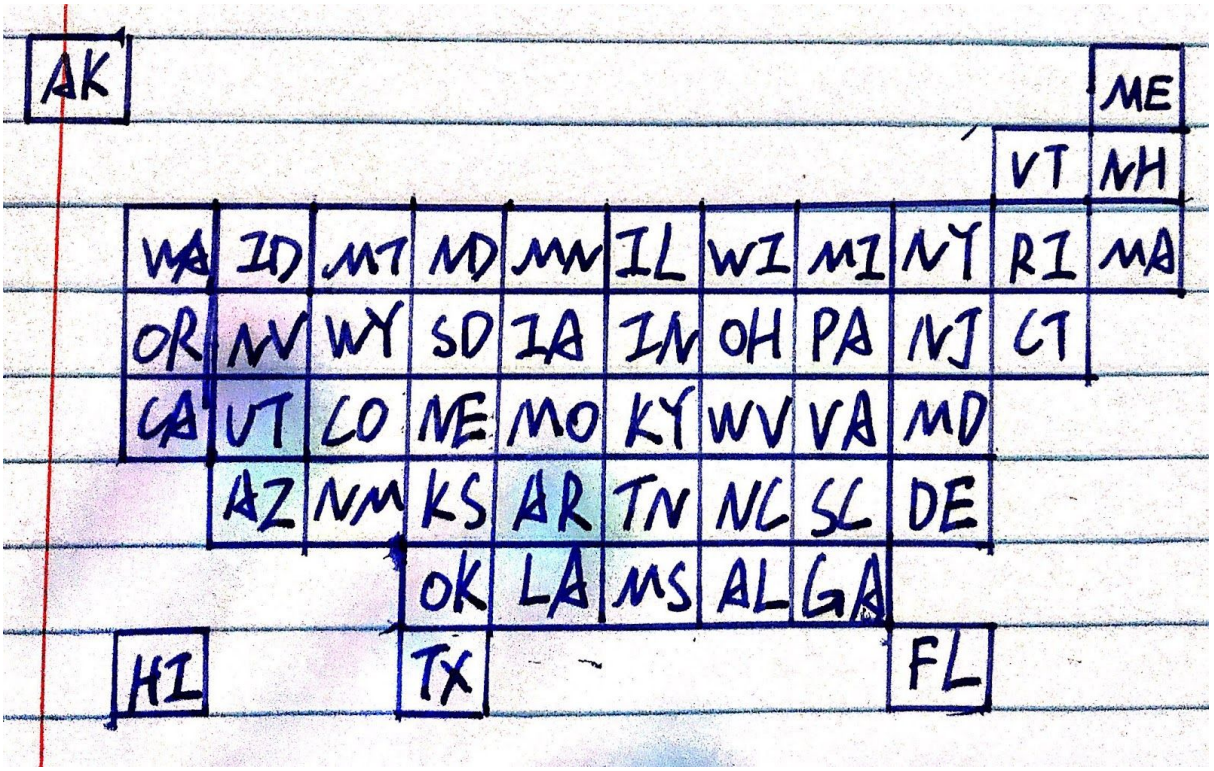
When the mouse hover on any one of the chart, the region for the corresponding year will be highlighted in all of the charts in these area. For example, if the mouse is within the range of 2008 on the National House Median Price chart, then the region for 2008 in all of the charts will be highlighted.

On the left, there will be a US map displayed, and major cities will be marked on this map. The map can be zoomed in and zoomed out. When the mouse hover on a city mark, a tool dip will be displayed which has the information of the median. You can also click on the mark of city. Click once on a city means this city is selected, click again on this city means this city unselected. You can select several cities.

Under the US map, there will be a chart for the Median price of houses of the selected cities during the chosen time window. The mark will be line, and the channel will be position. For different cities, different color will be used.

Design 2:

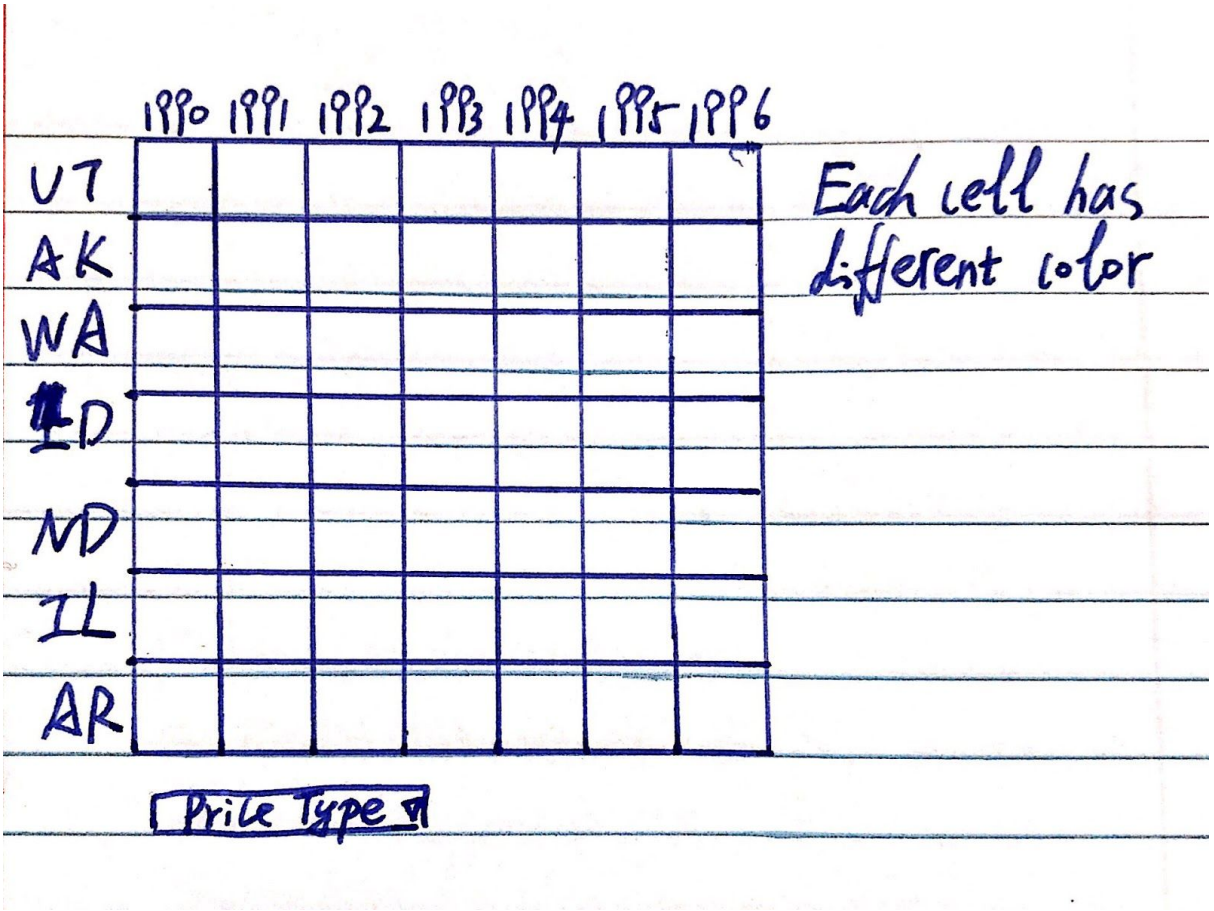




In this design, the main line chart in the center and control panel is on the right.
On the control panel, there is a state tiles of U.S. When click on the state, a drop-down menu is displayed to let user to choose one city from the chosen state. Below the state tiles, there is a time line display. A brush is designed on the time line chart to select the time period.

On the main area, there will only one line chart. Based on the selection of control panel, different line will be display in this area. The purpose of single chart is trying to make the display simple enough that help users focus on the data analysis. A drop-down menu will be placed above the line chart, which is used to indicate which price should be displayed.

Design 3:



In this design, instead of using the lines to represents the trend of housing price, we decide to use the color to embed the price information.
At same time, we also need the U.S tile chart in design 2 to choose the regions that needed to be compared.

Each cell in the table have a different color based on chosen price type.
Furthermore, when use click on one of the cells, the price of that cell will be the benchmark and all the color of other cells will change based on this.
This design will show us the what's the difference between the chosen cell and all other cells.

Must-Have Features:

1. The comparison between different locations through the time.
2. How does the housing prices change over all these years.
3. Show the connections between different factors and the housing prices.

Optional Features:

1. Display the derived data of house price, like the rate of rise.

Project Schedule:

1: Week 1 (Nov 4th):Discuss with TA to decide which design we are going to use. Have all of the data ready.
2: Week 2 (Nov 11th): The comparison between different locations through the time.
3: Week 3 (Nov 18th): How does the housing prices change over all these years.
4: Week 4 (Nov 25th): Show the connections between different factors and the housing prices.
5: Week 5 (Dec 2nd): Try to work on the optional feature(s).

Feedback from peer review:

Group members reviewed our plan: **Shenruoyang Na** and **Di Wang**

The comments from Na and Wang are very helpful.

Suggestion 1:

Add a selection feature for the map.

How to address:

We will add a selection tool for the map we are going to display. This selection tool will be like a brush. With this selection tool, we can select one or more cities on the map, and show the house price/ rental price for each selected cities in a new chart.

Suggestion 2:

Find a way to visualize the geometric and population data for different cities.

How to address:

We add a tooltip. When the mouse hover over the city, information about this city, including which state it belongs to, population, and area will be displayed.

Suggestion 3:

Using hue of color as a channel to visualize the change of median house price of different states during a time period on a map may not be the best choice. Instead, use line chart might be a better choice.

How to address:

Agree. Map is useful for display geometric data. Since we want to visualize change of house of price of different states, position should be a better channel for this purpose.

Suggestion 4:

The basic features of this project are not very innovative, because there is already different visualization of house prices available on line, for example, Zillow. But the optional feature, including visualizing the derived data of house price (e.g. rate of price change), correlation of other factors with house price (e.g. unemployment rate, mortgage interest rate), and prediction of house price are interesting, which will lead to an interesting story.

How to address:

We are going to finish these optional features as many as possible.

2017-11-7

Meeting: Discuss who will be responsible for what. Yizhu will be responsible for data processing, framework setup, and map part of the visualization. Han will be responsible for the line chart part and the extra features.

2017-11-8

1. Collect all the data we need and transform the data into json file with Python script.
2. Set up the basic framework with bootstrap.
3. Collect the map data from: http://eric.clst.org/Stuff/USGeoJSON

2017-11-13

1. Setup the basic div structures and css style.

2017-11-14

Meeting: Discuss about the detailed design, including the style of the map, the styple of the check boxes, and the communication between the map and the line chart.

2017-11-15

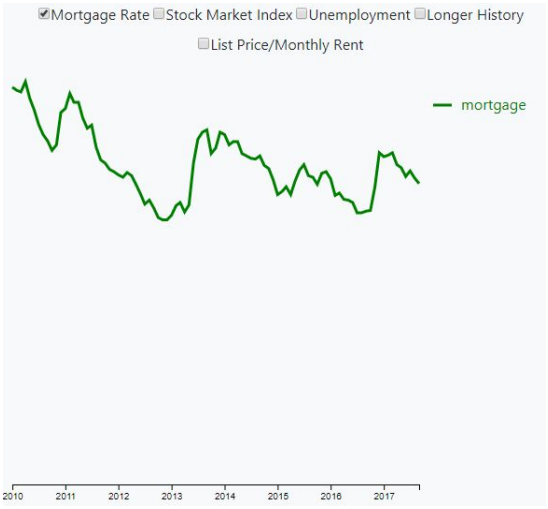
1. Add the check boxes.
2. Draw the line chart with dummy data.

2017-11-17

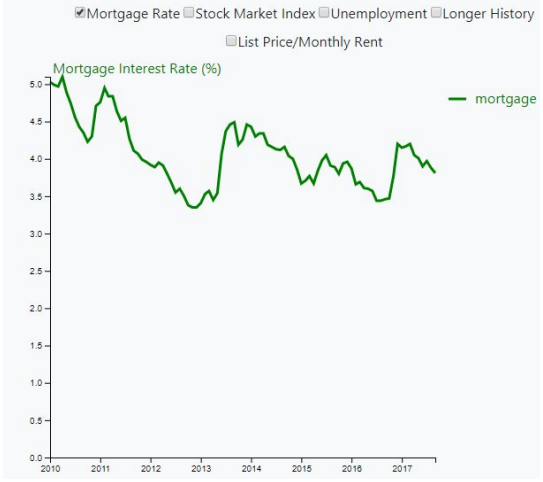
1. Draw the economy line chart with the real data.

2. Since the real data contains date which is not the common format recognized by javascript. We have to convert the date into right format, which took us some time to figure it out.
3. Add new feature: Since the economy chart contains lines of data with different unit, we decided not to show y-axis until the mouse move over on the line.

Before the mouse move over the line, no y-axis displayed:

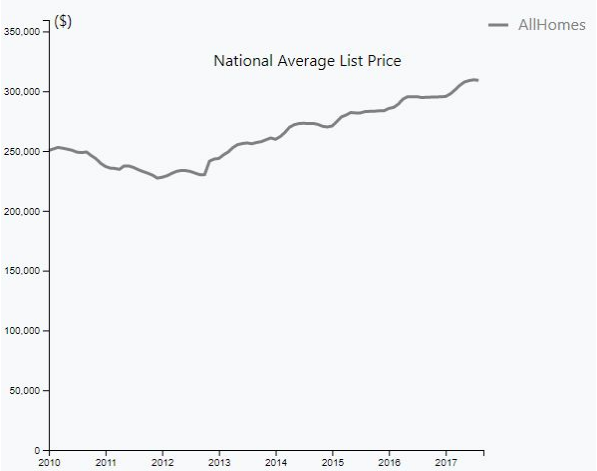


When the mouse move over the line, y-axis was displayed:



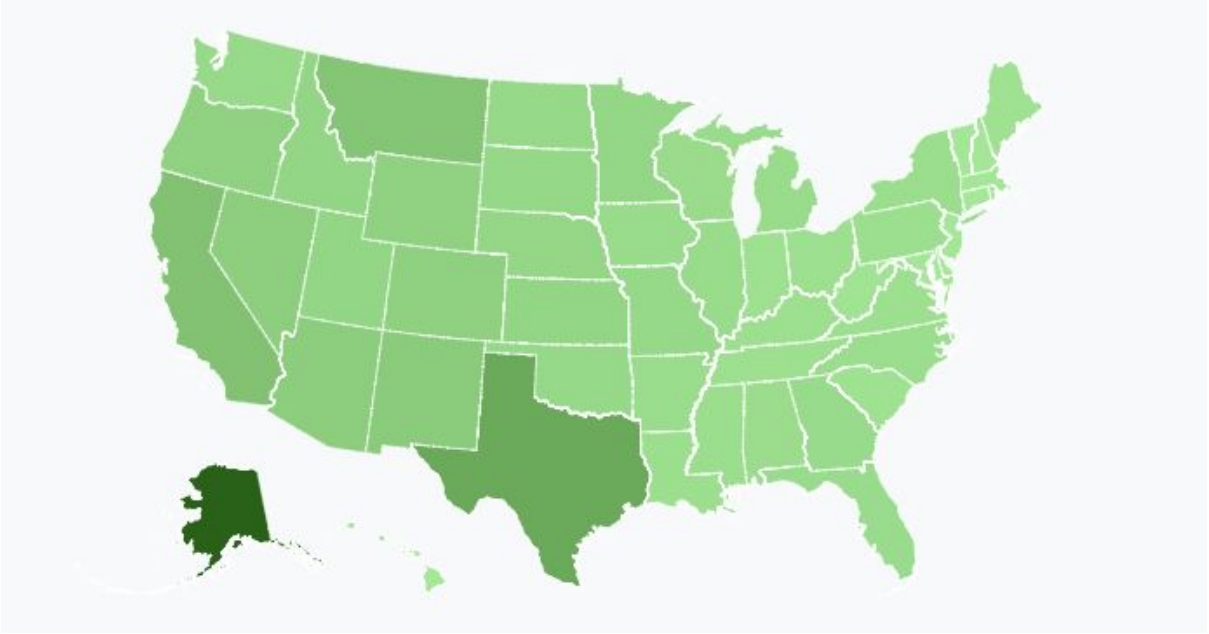
2017-11-19

1. Draw the price line chart with the real data.
2. Only the line for the national average prices were drew at this stage. Since the original data do not have the national average prices, we have to build this feature with derived data.



2017-11-20

Draw the basic USA map. This is easy part. Now I have the map like this:



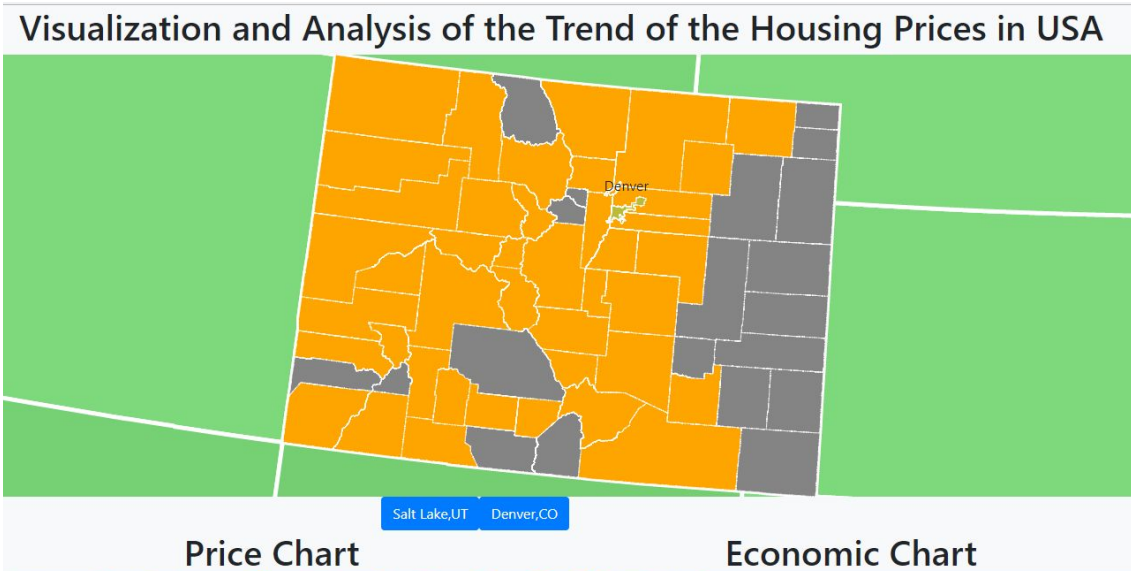
However, the difficult part is when the user select a state, how do we display the map of the chosen state and after that how can we go back the overall map?

Difficult I: Go inside into the specific state

This problem troubled us for long time. We could not figure out how to go inside a specific state.
We searched for this problem but it seemed no one has encountered this.

So, we decided solve this by ourselves. This is the center part of our project and we did not want to give up this.

The final solution of this problem is we are using two layers of map.
The default layer is just the overall USA map and the second one is the map of chosen state.
At first, when we load the data, we only make the first layer visible.
When user click on a specific state, we make the second layer visible which covers the first layer to show the county map of the chosen state.



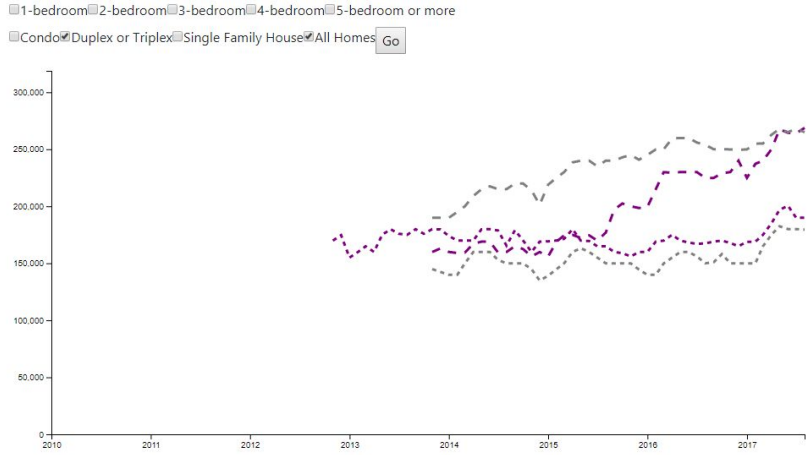
2017-11-21

Meeting: Discuss about the set the width of the svg to be ajustable

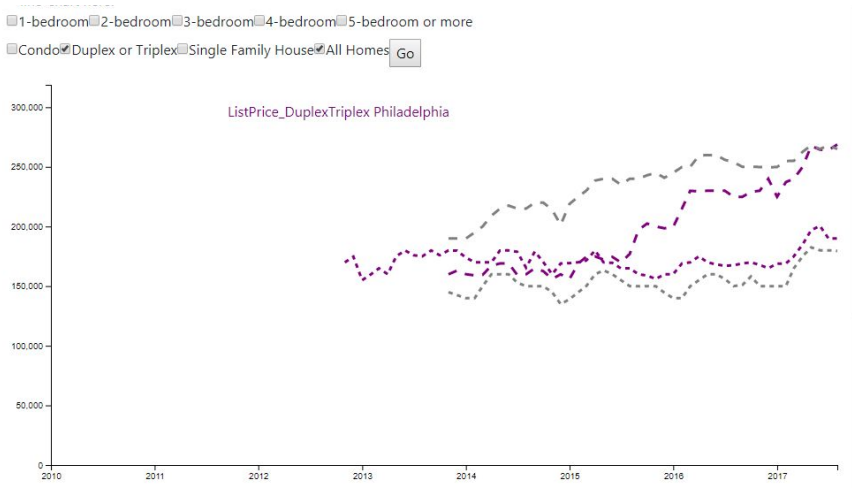
2017-11-23

1. Finish the price line chart
2. To Since there might be more than one lines displayed on the chart at the same time, the name of a line will be displayed when the mouse move over the line.

Before the mouse move over the line:



When the mouse move over the line of List Price of Duplex in Philadelphis



3. Push the code for drawing the line charts to the git.
4. Now, we have all of the necessary features.

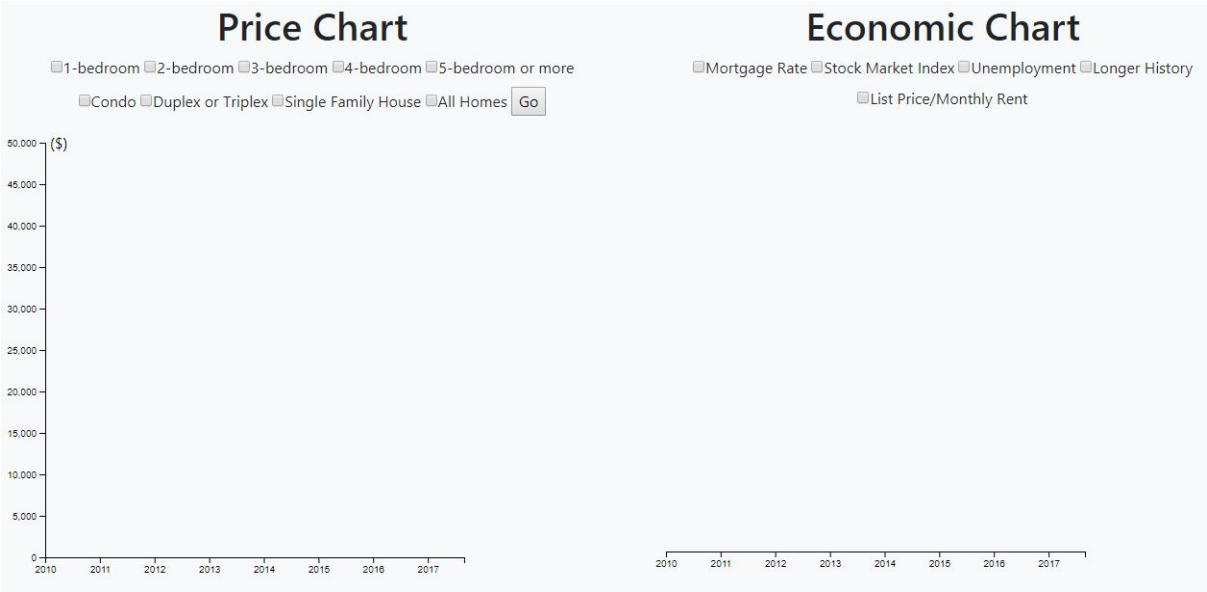
2017-11-25

After solving the state map problem, the next problem is how to draw these two layer maps using the same projection. The simple solution to this is using the translate function. But this seemed to be problematic because this will affect the projection result. I need to calculate the offset manually of each state which is impossible for us.

Luckily, when I read about the basic concepts of SVG, I learned the concept of viewBox. My solution for our problem is trying to change the viewBox instead of using translate function. The advantage of viewBox is that it will not affect our projection method and provide a smooth animation, which is very cool.

2017-11-26

1 meeting: Discuss about adding the following feature: a. Display the x-axis all of the time. b: Add legend for the line chart
2
Add a new feature: display the x-axis for the line chart all of the time. Previously, the x-axis was only displayed when specific data is selected. To make it look nicely, we decided to display the x-axis all of the time.



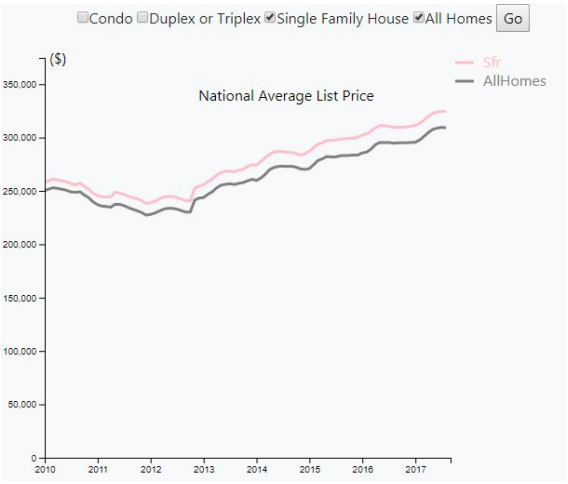
2017-11-27

1. Fix a bug to display the county name. When show the county name, we assign an id to the text of the county name. However, there are some counties, like “Salt Lake” contains space in the name, which can not be used as an id.
2. Add a new feature: adjustable width for the line chart

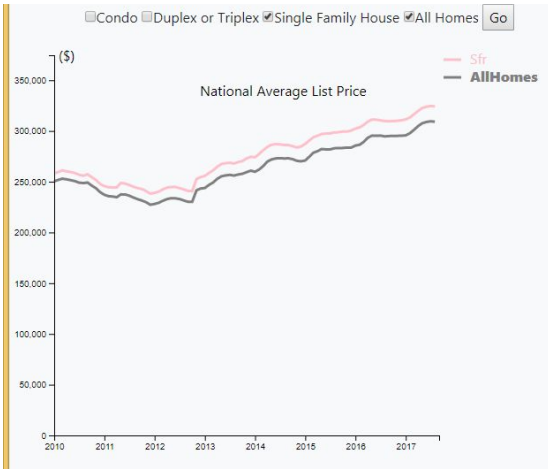
2017-11-28

1. Add a new feature: display legend for the price chart. When the mouse move over a line, the corresponding legend’s text will become bold.

Before the mouse move over the line:

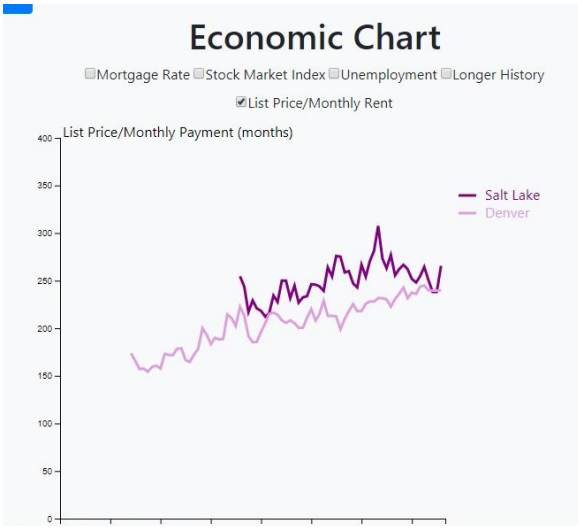


When the mouse move over the line of all homes, the text of the legend of all homes will become bold.



2017-11-29

1. Add the option feature: display the list price/monthly rent ratio for different counties. Here, we use derived data.



To calculate this derived data, we need the average list price and the monthly rent. However, in the zillow dataset, one or both of these two data were missed for some counties. So, we can not get this information for all of the counties of US

2017-12-01

1. Add a new feature: display legend for the economic chart.

2017-12-02

Record the video