MSDS622 Visualization Project Proposal

● Basic Information

Project Title: An Interactive View of US Housing Price Index Changes

Team Members: Bingyi Li, Guoqiang Liang

E-mail Addresses: bli36@dons.usfca.edu, gliang5@dons.usfca.edu

Github Repository: https://github.com/asenliang/MSDS622_Visualization

Website: https://asenliang.github.io

● Background and Motivation

As master's candidates at University of San Francisco, we are expected to become great data scientists after graduating from the Data Science Program. It is likely that we will have high salaries, stocks and shares in the near future, and will be able to purchase the first apartment, condo or even a house in our life. So it is important for us to know the housing price in the United States. What does it look like now and how did it change throughout the history? In order to learn these information, a deep understanding of Housing Price Index (HPI) will be necessary.

As explained by Federal Housing Finance Agency, "The HPI is a measure designed to capture changes in the value of single-family houses in the U.S. as a whole, in various regions and in smaller areas. It serves as a timely, accurate indicator of house price trends at various geographic levels. It also provides housing economists with an analytical tool that is useful for estimating changes in the rates of mortgage defaults, prepayments and housing affordability in specific geographic areas. "

There are lots of HPI-related data available online, but it is hard to learn directly from those bunch of numbers. So we decide to create a website to better visualize the information.

● Project Objectives

The primary goal of this project is to create an interactive visualization of historical U.S. Housing Price Index values. More specifically, we identified a few detailed objectives:

1. Compare the HPI of different states at a historical time point
2. Visualize changes of HPI of a state over time
3. Provide more detailed HPI of different counties within a state

From our visualization, users will gain deep understanding of housing price changes in the United States in general, and more specifically, changes within different states and counties.

● Data

We have picked up four HPI datasets to help with the analysis from the official website of Federal Housing Finance Agency.

1. Quarterly U.S. HPI summary

   https://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_AT_us_and_census.csv

Country	Year	Quarter	HPI(1980 as base)
DV_ENC	1975	1	63.6

2. Quarterly HPI by state

   https://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_AT_state.csv

State	Year	Quarter	HPI(1980 as base)
AK	1975	1	61.65

3. Annual HPI by county

   https://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_AT_BDL_county.xlsx

State	County	FIPS code	Year	Annual Change (%)	HPI	HPI with 1990 base	HPI with 2000 base
AL	Autauga	01001	1986	.	100.00	95.40	71.03

4. Annual HPI by zip-code

   https://www.fhfa.gov/DataTools/Downloads/Documents/HPI/HPI_AT_BDL_ZIP5.xlsx

Five-Digit ZIP Code	Year	Annual Change (%)	HPI	HPI with 1990 base	HPI with 2000 base
01001	1985	.	100.00	62.26	61.60

- **Data Processing**

  Although the columns and features in these datasets are all slightly different from each other, they all include the information of HPI values over a similar time span, from 1970s to present. So we can join those data and extract information without performing substantial data cleanup.

  However, it should be noted that the HPI values in these datasets are not necessarily having the same basis. For example, some datasets build on the assumption that the HPI values in 1980 are always 100, while this date might change in other cases. Therefore, we are more interested in the relative ratios of HPI rather than their absolute values, which means we need to line up the dates and recalculate the indices when necessary.

- **Visualization Design**

  To start with, we will provide a geo-map of the United States, where each state is colored base on its yearly HPI. Users can change timestamps to get different geo-maps for different years. Users can also click on each state to see the HPI changes within that state. (Appendix: Graph 1)

  Once users click on a state, the diagram will zoom in and show a geo-map of that state, where each county is colored based on its yearly HPI. A time series diagram will show up next to the geo-map, demonstrating the quarterly change in HPI of that state and of the country. So users can know how does the housing price in that state compare to national housing price. (Appendix: Graph 2)

  Users can also choose a county, possibly by clicking on it or choosing from a list. Then a new line will be added to the time series diagram, representing the HPI changes of that county. (Appendix: Graph 3)

● Must-Have Features

1. A geo-map of the US, where each state is clickable and colored base on its yearly HPI.

2. Geo-maps of each state and county, where each of them is clickable and colored based on its yearly HPI.

3. Time series diagrams for each state and county.

● Optional Features

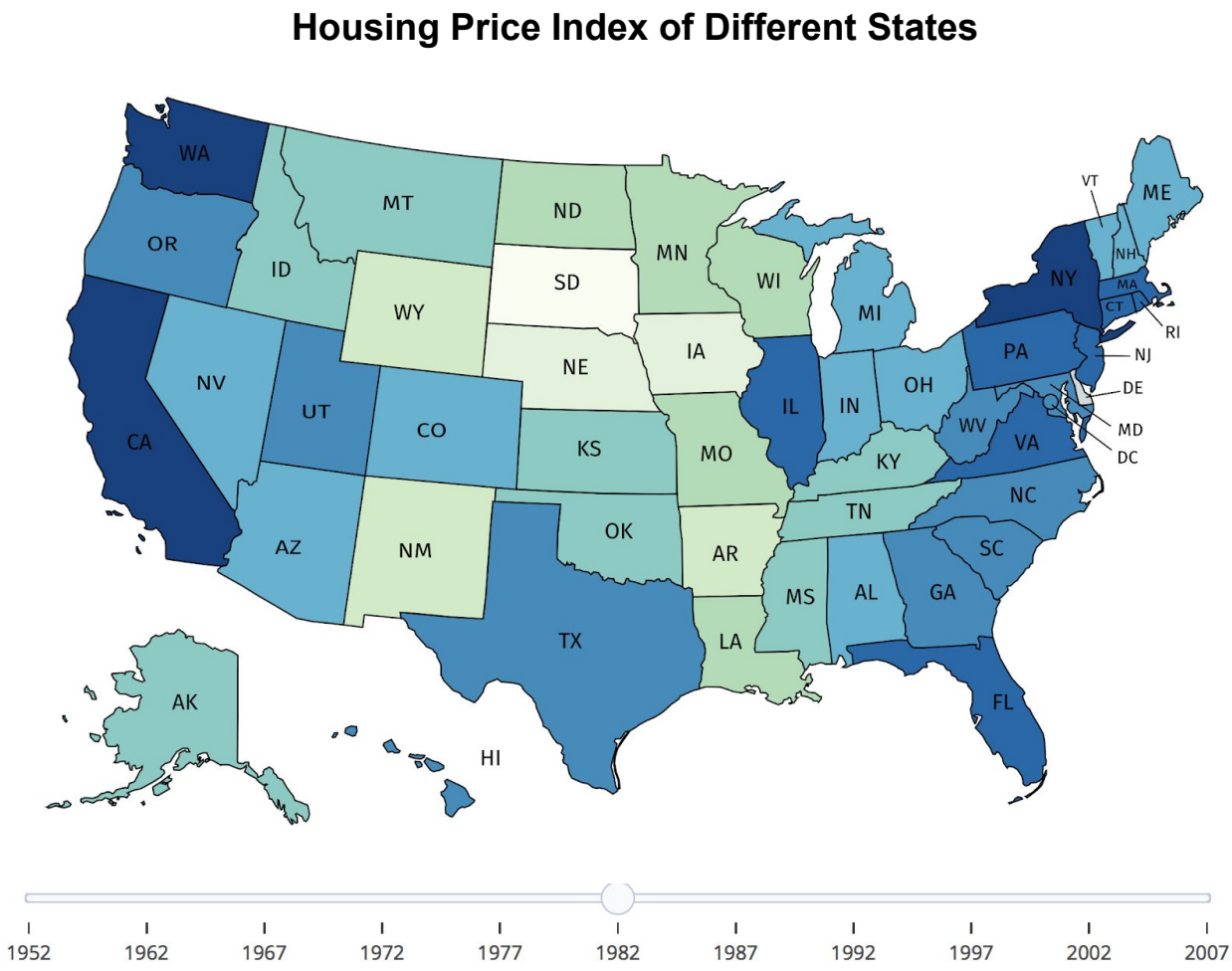
In addition to the features described in the section above, we could potentially add a search box such that users can retrieve the HPI data of a given zip-code, and a geo-map comparing HPI of the adjacent regions would be provided. Also, since we are visualizing time series datasets, we would like to include a sliding window (similar to the one for geo-map) in the line charts, and users could choose the date range they are most interested in. (Appendix: Graph 3, 4)

● Project Schedule

Package: *Plotly;* **Work Style: *Pair Programming***

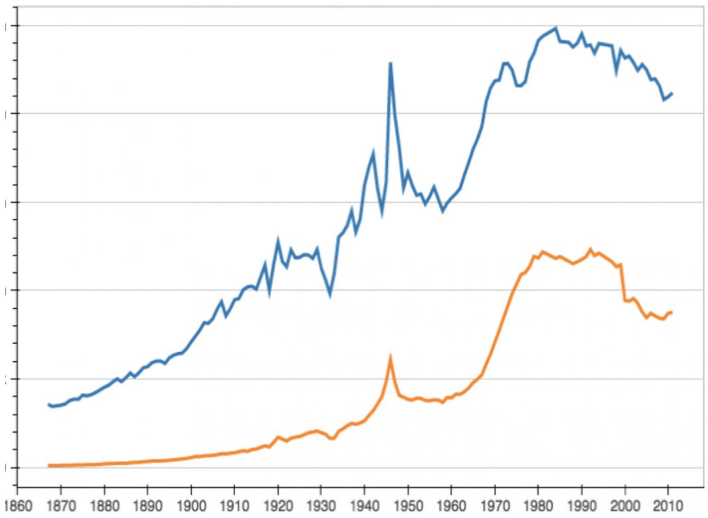
Week of ...	Schedule
04/16/2018	1. Write project proposal 2. Design a website prototype 3. Generate the country geo-map in Jupyter notebook
04/23/2018	1. Implement the interactive geo-map on the website 2. Implement the interactive state map and time series diagram
04/30/2018	1. Implement the interactive county map and time series diagram 2. Finalize the website design
05/07/2018	1. Prepare for the presentation 2. Start to work on optional features
05/14/2018	1. Attempt to implement the zip-code feature 2. Peer review

Appendix: Graph 1



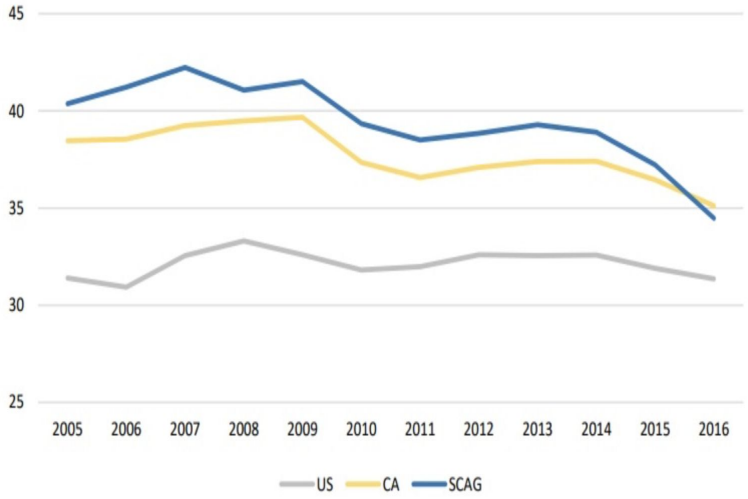
Appendix: Graph 2

Housing Price Index of California



Appendix: Graph 3

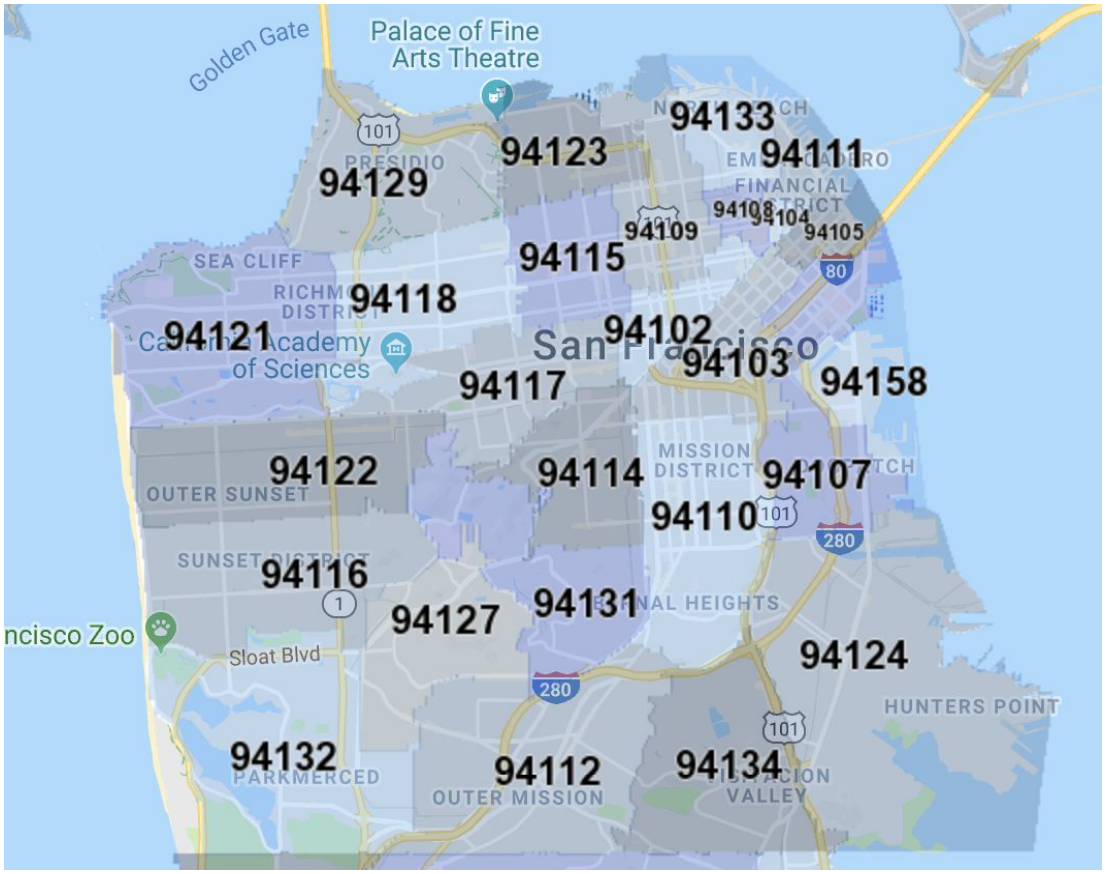
Housing Price Index of XX County, CA



Appendix: Graph 4

Housing Price Index of xxxxx Zip Code

Zip Code: 94117



1952 1962 1967 1972 1977 1982 1987 1992 1997 2002 2007