**Technosapiens**

House Price Prediction

**Members -**

Soumyendra Shrivastava (016670121)

Tanuja Reddy Maligireddy (016715348)

Siddhant Sancheti (016710421)

# **Objective**

Building a model that can precisely forecast house prices based on a collection of input data is the goal of a house price prediction project. This can assist developers, investors, and real estate agents in making educated decisions on the value of the real estate, as well as house buyers and sellers.

1.1. **Business Logic**

In order to increase the efficacy and efficiency of real estate transactions and assist informed decision-making in the housing market, the business logic behind home price prediction is founded on the notion that reliable and timely information on future house prices may be provided. All parties involved may benefit from this, and the housing market may expand and develop as a result.

# **Goals**

1. The most typical application of house price prediction is to determine a home's current market value. It can be helpful for both homeowners who want to estimate the value of their home and prospective buyers who want to assess whether a home is priced reasonably.
2. Assessing the possible return on investment (ROI) of a real estate investment: Investors can use home price forecasts to assess the possible ROI of a property and decide whether to buy or sell in accordance with the results.
3. Increasing the accuracy of real estate evaluations: Real estate appraisals frequently depend on judgments about how much a property is worth. By supplying unbiased data, house price prediction algorithms can increase the accuracy of appraisals.
4. Finding locations with the greatest potential for price growth: By forecasting house values in various communities or cities, developers, investors, and governmental bodies can choose locations with the greatest potential for price growth and focus their investments there.
5. Real estate market efficiency can be increased by minimizing the time and effort needed for buyers and sellers to come to an agreement on a price. Accurate house price predictions can do this.
6. A house price prediction project's overall objective is to offer knowledge and data that can assist people and businesses in making wise real estate investment and transaction decisions.

# **Data Narrative**

Our Datasets contain the following information -

1. **FHFA Dataset**

The Federal Housing Finance Agency (FHFA) serves as a source of data and research on the U.S. housing finance system and provides information and analysis on topics such as housing prices, mortgage rates, and market trends. This information can be used by a variety of stakeholders, including policymakers, industry professionals, and consumers, to make informed decisions about the housing market and the future of mortgage finance in the United States.

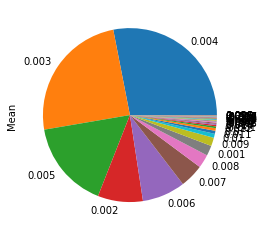
A House Price Index (HPI) is a statistical measure that tracks changes in the prices of residential properties over time. It can be used to measure changes in the overall market and changes in specific regions, cities, or neighborhoods. The index is calculated based on data collected on sales of individual homes, and it typically reflects the change in prices for a representative basket of properties in the market.

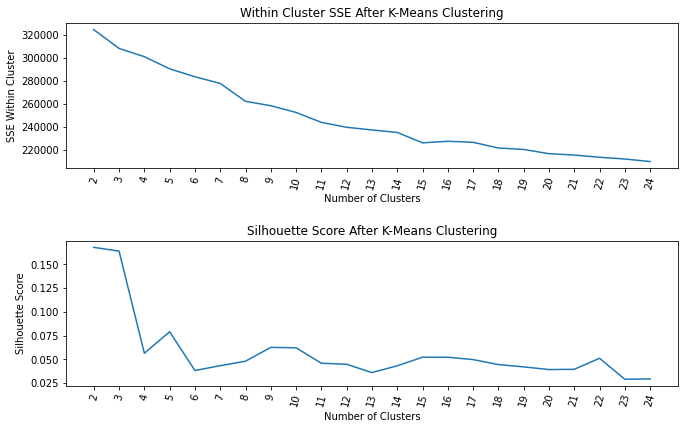
In a house price prediction project, the HPI can be used as a benchmark for evaluating the performance of the prediction model.

1. **Zillow House Price Dataset**

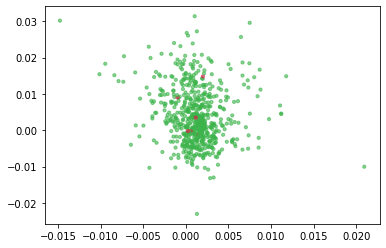
The Zillow House Price dataset includes a substantial amount of data on a variety of housing market variables, such as median home values, median rent costs, and the availability of homes for sale in a particular area. The information is gathered at the zip code, city, and state levels and includes a sizable number of US urban regions.

# **Project Flow**

1. Maximize ROI:
   1. Preprocessing
      1. Mean of Month on Month Change 
   2. Using Fractal Clustering with K-Means:

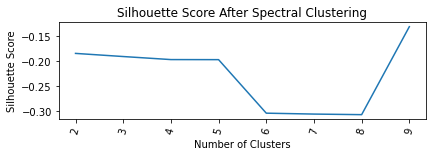


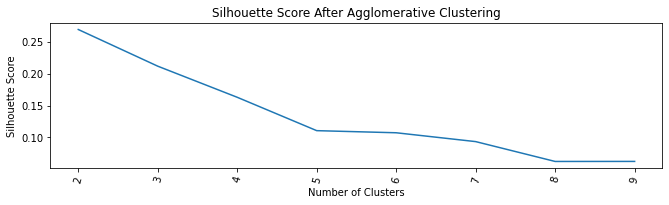
* + 1. Golden Cluster

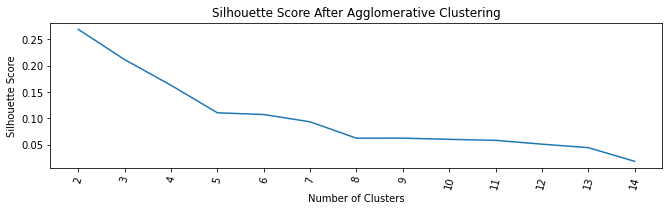


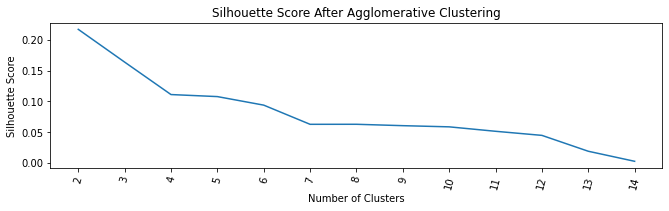
### Finding Most Suitable Algorithm



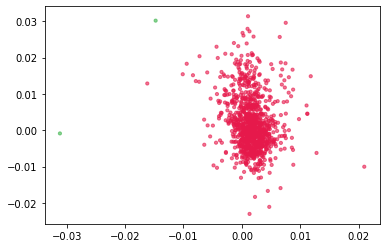




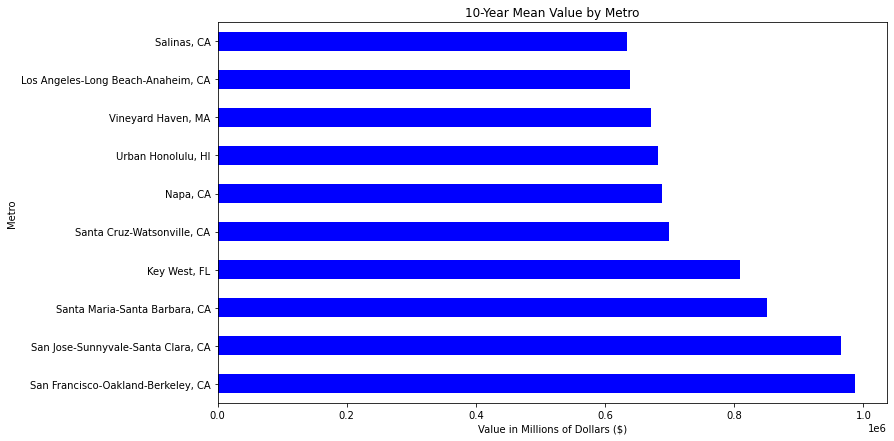
* 1. Applying Fractal Clustering with - Agglomerative Clustering
     1. First Iteration
     2. Second Iteration

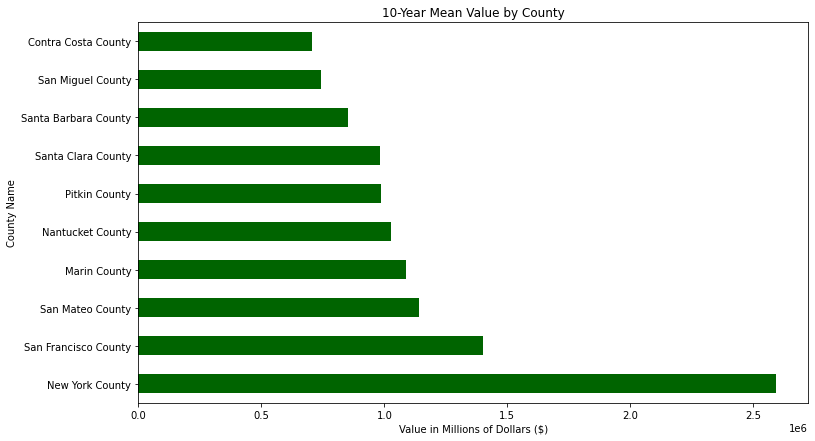


* + 1. Golden Cluster

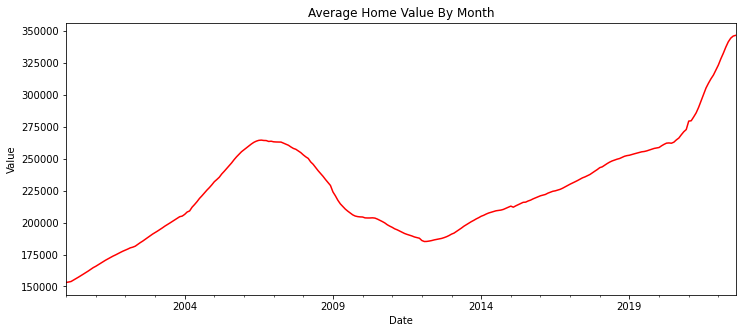


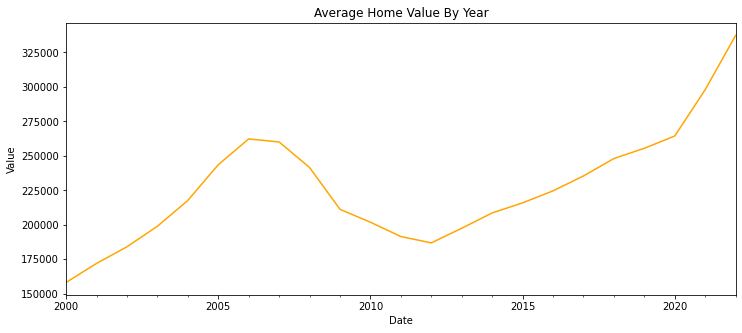
1. Finding the Best Location
2. Exploratory Data Analysis:
3. Top 10 County and Metro(Mean)





1. On Zipcodes:



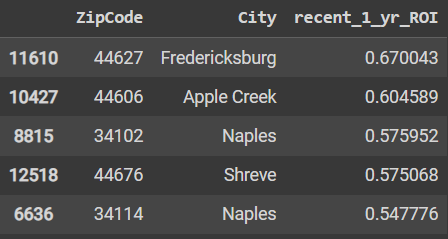


1. Lowest value of recent one-year ROI:

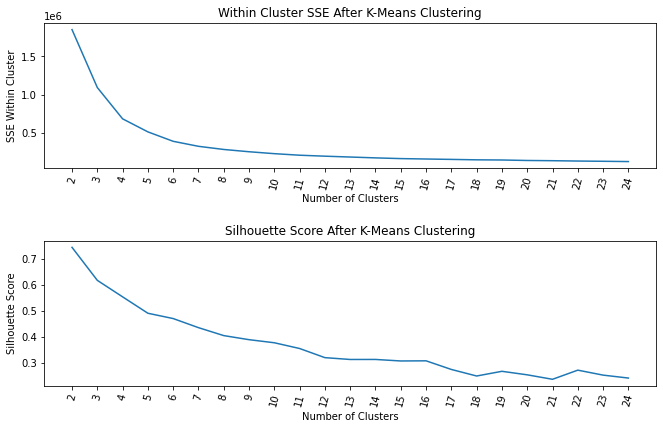


1. Highest values of recent one-year ROI:

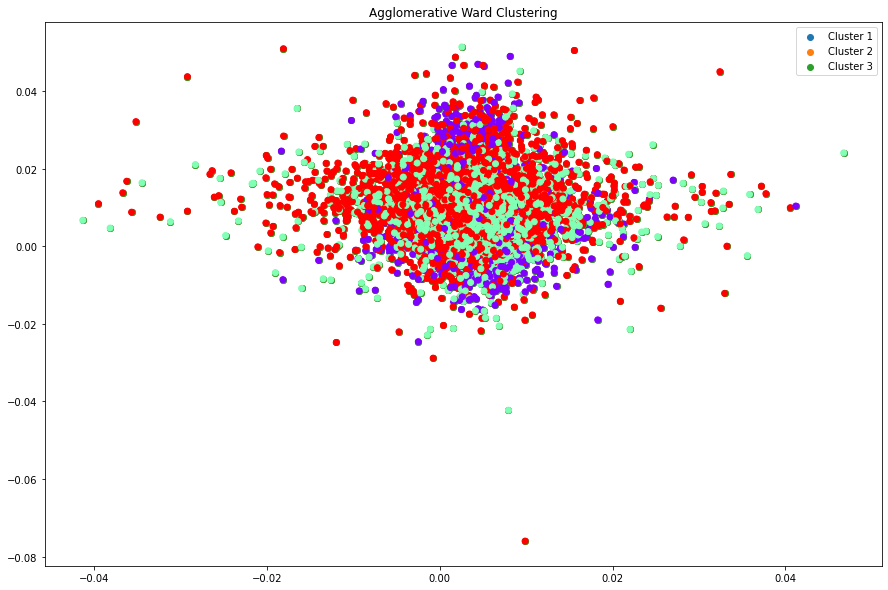
Below we can see the top cities with zip codes having the highest recent ROI



1. Applying Clustering Algorithms:
2. K-Means Clustering:

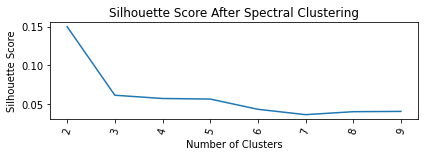


1. Agglomerative Clustering:



silhouette score: 0.07

1. DBSCAN
2. Spectral Clustering:



Conclusion: Comparing silhouette scores for all the algorithms we can conclude that Agglomerative fits the best for our use case.

1. Fractal Clustering with K-Means:
2. First Iteration:

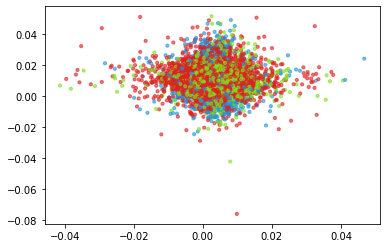
cluster cluster

0 0 5872

1 1 3431

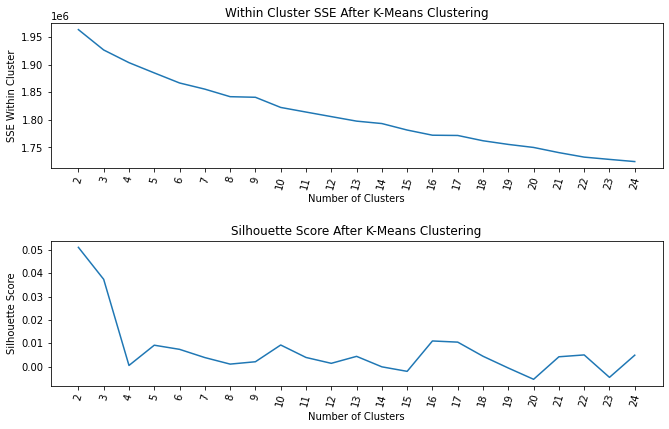
2 2 2298

Name: cluster, dtype: int64



From this, it’s clear that Cluster 0 can be primary cluster for 2nd Iteration

1. Second Iteration:



From above graph, we know that the number of **Optimal Cluster is 3.**

**cluster cluster**

**0 0 831**

**1 1 3441**

**2 2 1600**

Cluster 2 seems to have more density to churn further through iteration third

1. Third Iteration:

cluster cluster

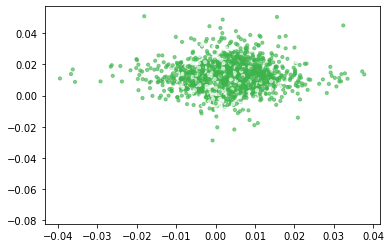
0 0 24

1 1 1739

2 2 1678

On performing the third iteration we found that 2 clusters are almost shared divided. Hence, one prominent cluster from **iteration 2 itself will become our golden cluster**

**Golden Cluster Plot:**

****