

Abstract:

As the pandemic just happened, the real estate market is booming, and it is facing different challenges. We will choose the finance side of the real-estate market. We will work on the Data set and API from Zillow.com.

Datasets:

Following is the link of the CSV files we will be working <https://www.zillow.com/research/data/>

* We will merge the dataset available by RegionID and get everything in the one CSV.
* https://github.com/alisaghaffi/project-1/tree/main/Resources

Research Question to Answer:

Housing analysis will be done on the general basis such as housing prices, inventory, sales count, availability and buyers. Some of the key points that we will be focusing on are the questions below.

* Are there significant differences between states in price ? ali-saghafi
* Are there significant differences between cities of California in prices ? ali-saghafi
* Are there any seasonal trends throughout the year with houses being sold or rented? -Dennis
* Changes in price over time by location, starting price, and size - Justin (Location)
* Compare the percent increase of house value vs house renting price -Dennis
* The comparison of rent prices by region - Luis
* An estimate of the current market value- fenny
* Highest and lowest reasonable price range- fenny
* Neighborhood property value average- fenny
* Map showing comparable sales data

If we include a dataset from Kaggle, we might be able to analyze additional questions such as below.

* Price per square foot
* Per bedrooms
* By location
* Does distance from school affect cost?
* Does distance from the airport affect cost?
* 1 story vs 2

Rough Breakdown Of Tasks:

* Clean data
* Merge data
* choose 1 or 2 question to look for answer too
* plotting graphs
* heatmaps?