Housing Data Project

Tristan Perry
Kelsea Thayer
Owen Pollard
Brittney Watts

Project Overview

Our goal was to conduct a housing market analysis to analyze market trends over time, particularly mean home values and rent prices for several cities across the United States over the span of 10 years(2015-2024). We used Zillow and census information to create databases for the rental prices and housing values over every month over the past ten years!

Why did we choose this data?

We chose this data because it was easily accessible and we believe everyone should have access to available housing information. It is also important for people to know market values to avoid discrimination when trying to secure housing. From this data, anyone can find changes in housing market values over time and analyze their trends!

Technology Used

- Jupyter Notebook
- VS Code
- Pandas
- SQLite
- Flask
- Excel
- Collections

Data Sources

We retrieved data from Zillow's "Housing Data" website for researchers. The datasets we retrieved are:

- "Zillow Home Value Index(ZHVI)" shows the values of homes within the 35th to 65th percentile for the given regions.
- "Zillow Observed Rent Index(ZORI)" shows the values of homes within the 40th to 60th percentile for the given regions.

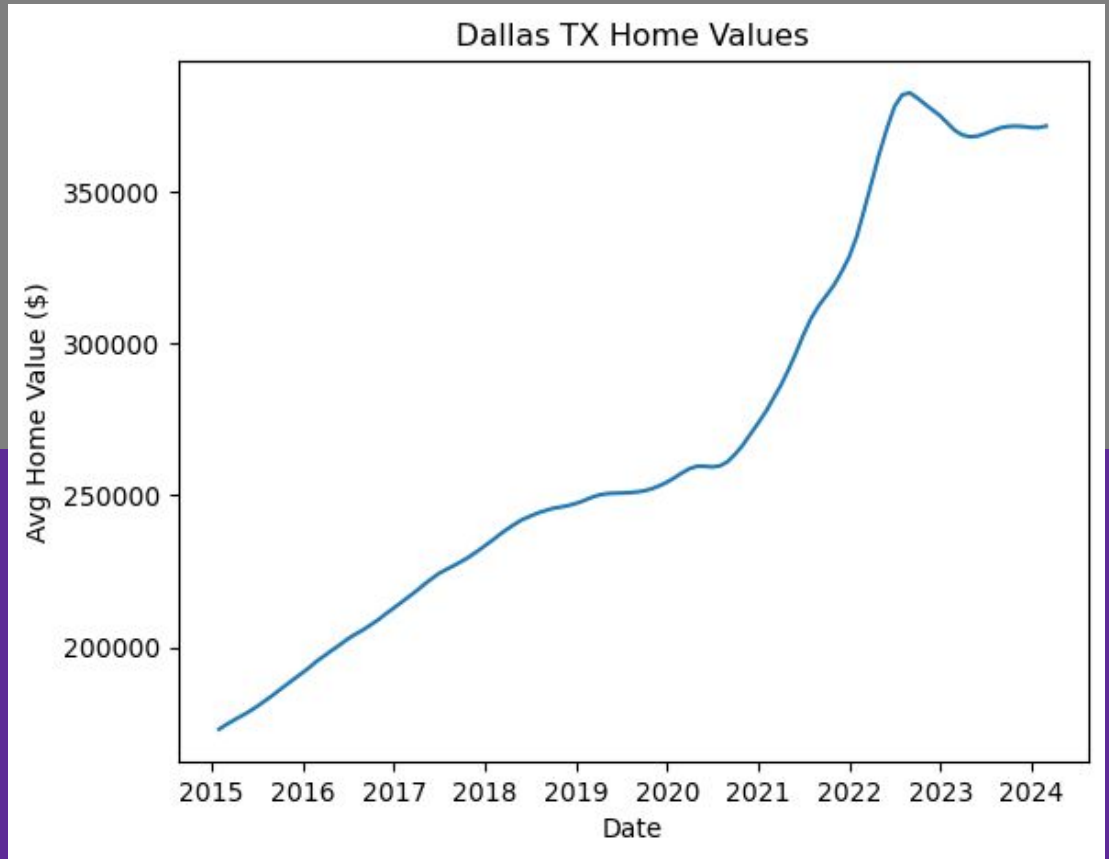https://www.zillow.com/research/data/

Rental Value Increases

Rental Values Comparison by State: February 2015 vs February 2024
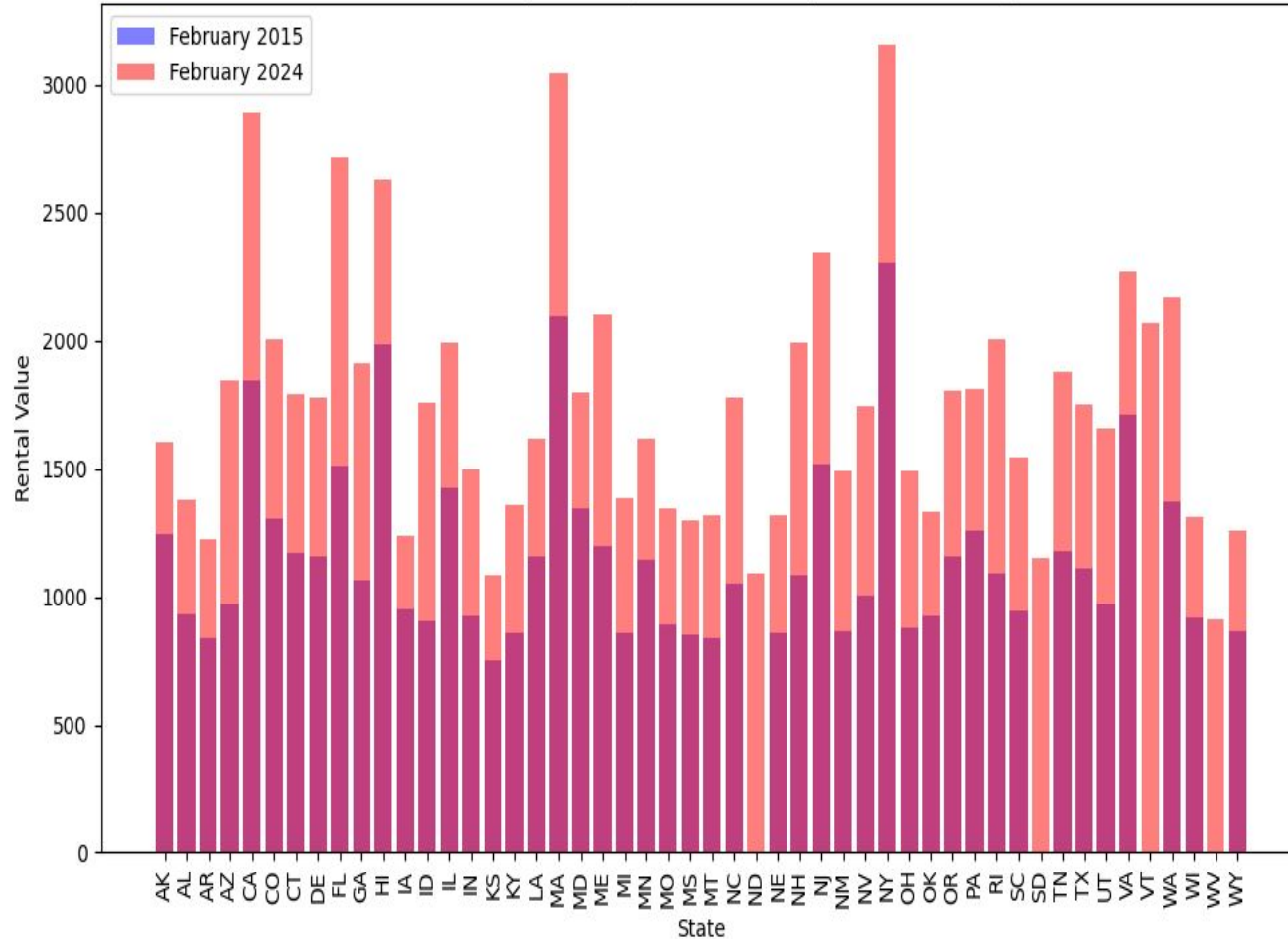
This is a graph of home values for Dallas, TX from 2015-2024. Most cities within the data follow the same general trend.

Home Values Over Time

Housing Value Statistics for the Top 5 Cities and Lowest 5 Cities

	Edwards, CO	Santa Cruz, CA	Jackson, WY	San Francisco, CA	Vineyard Haven, MA	Helena, AR	Murray, KY	Coffeyville, KS	Forrest City, AR	Danville, IL
count	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00
mean	817914.57	895937.61	901242.90	957895.87	973687.66	51553.91	60556.23	64905.90	65387.77	67531.58
std	232554.42	176577.51	281864.24	154877.04	271159.27	3502.19	20446.61	9510.11	4859.85	6778.71
min	553217.17	626403.23	556253.00	672632.12	662373.53	43735.44	36254.23	55272.99	55342.53	59348.20
25%	643930.60	763005.00	683987.96	826266.52	747012.92	49232.75	40805.20	57027.95	62056.37	62998.21
50%	712566.85	838462.29	783098.82	941325.76	874334.54	52358.79	52732.18	61097.86	66738.90	63798.75
75%	1009071.43	1080414.64	1165321.90	1114335.16	1201604.18	53772.28	81600.85	73629.26	69867.39	71915.64
max	1266339.26	1237211.61	1441041.16	1259653.45	1465742.97	58560.70	97566.26	82443.67	72530.44	82741.36

Top 5 cities with the highest average home values:
1. Vineyard Haven, MA
2. San Francisco, CA
3. Jackson, WY
4. Santa Cruz, CA
5. Edwards, CO

Retrieved from Zillow data that reports the values of the 35th to 65th percentile of all homes in the US. All values were found by sorting the mean values for every city, and extracting the 5 cities with the highest and lowest values.

Bottom 5 cities with the lowest average home values:
1. Helena, AR
2. Murray, KY
3. Coffeyville, KS
4. Forrest City, AR
5. Danville, IL

Rental Cost Statistics for the Top 5 Cities and Lowest 5 Cities

	Oxnard, CA	Los Angeles, CA	Santa Maria, CA	Santa Cruz, CA	San Jose, CA	Evansville, IN	Canton, OH	Lawton, OK	Topeka, KS	Davenport, IA
count	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00
mean	2334.10	2351.98	2530.81	2747.96	2924.83	669.25	763.15	801.44	805.60	818.93
std	350.04	317.85	544.18	429.03	187.21	91.56	118.02	116.49	113.96	77.82
min	1829.17	1833.21	1800.52	1973.74	2503.95	553.26	629.92	691.35	669.11	719.76
25%	2056.51	2119.95	2103.88	2429.12	2773.65	594.67	667.67	712.25	718.34	759.35
50%	2249.08	2334.30	2351.88	2692.00	2932.40	644.32	718.05	742.68	773.06	794.09
75%	2682.09	2621.07	3099.10	3112.59	3021.41	745.81	849.28	868.98	883.84	884.05
max	2956.98	2903.10	3527.31	3596.08	3253.34	879.17	1030.90	1049.82	1044.02	970.90

Top 5 cities with the highest average home values:
1.   San Jose, CA
2.   Santa Cruz, CA
3.   Santa Maria, CA
4.   Los Angeles, CA
5.   Oxnard, CA

Retrieved from Zillow data that reports the values of the 40th to 60th percentile of all homes in the US. All values were found by sorting the mean values for every city, and extracting the 5 cities with the highest and lowest values.

Bottom 5 cities with the lowest average home values:
1.   Evansville, IN
2.   Canton, OH
3.   Lawton, OK
4.   Topeka, KS
5.   Davenport, IA

Months with the Highest Rental Rates Increases

The highest rental rate spikes were in the summer of 2021 as people began moving again after a year of Covid. 53.3% of cities had their highest rental increases from June 2021-September 2021.

2021-06-30: 28 cities (14.5%)
2021-05-31: 25 cities (13%)
2021-07-31: 23 cities (11.9%)
2021-08-31: 21 cities (10.8%)
2021-09-30: 06 cities (3.1%)

- 'Los Angeles, CA': '2021-08-31',
- 'Chicago, IL': '2021-05-31',
- 'Dallas, TX': '2021-06-30',
- 'Houston, TX': '2021-06-30',
- 'Washington, DC': '2021-06-30',
- 'Philadelphia, PA': '2021-06-30',
- 'Atlanta, GA': '2021-08-31',
- 'Phoenix, AZ': '2021-06-30',
- 'Riverside, CA': '2021-07-31',
- 'Detroit, MI': '2021-05-31'

This is a snapshot of dates where the percent change of rent prices over all of the months from 2015-2024 is highest. There are 193 cities in this dataset.

Months with the Highest Home Value Increases

The housing prices spiked a year later in the Spring of 2022 as many people bought houses while interest rates were low. 59.6% of cities had their highest home value increases in the following months:

2022-03-31: 226 cities (26.8%)
2022-06-30: 108 cities (12.8%)
2020-11-30: 74 cities (8.8%)
2021-07-31: 64 cities (7.6%)
2022-04-30: 30 cities (3.6%)

- 'New York, NY': '2020-12-31',
- 'Los Angeles, CA': '2017-03-31',
- 'Chicago, IL': '2022-03-31',
- 'Dallas, TX': '2022-03-31',
- 'Houston, TX': '2022-03-31',
- 'Washington, DC': '2022-03-31',
- 'Philadelphia, PA': '2022-03-31',
- 'Miami, FL': '2022-04-30',
- 'Atlanta, GA': '2022-03-31',
- 'Boston, MA': '2022-03-31'

This is a snapshot of dates where the percent change of home values over all of the months from 2015-2024 is highest. There are 843 cities in this dataset.

Flask API

Available Routes:
The follwoing route takes input from the user, where City_ST can be replaced with a US city and state (state abbreviated). JSON format is returned for the maximum and minimum home values & date.

/analyze_hv?column_name=City_ST
Try: /analyze_hv?column_name=Dallas_TX for Dallas, Texas.

The follwoing route takes input from the user, where City_ST can be replaced with a US city and state (state abbreviated). JSON format is returned for the maximum and minimum rental costs & date.

/analyze_rc?column_name=City_ST
Try: /analyze_rc?column_name=Dallas_TX for Dallas, Texas.

127.0.0.1:5000//analyze_hv?column_name=Dallas_TX

```json
{
    "column_name": "Dallas_TX",
    "max_date": "Wed, 31 Aug 2022 00:00:00 GMT",
    "max_value": 382698.4094,
    "min_date": "Sat, 31 Jan 2015 00:00:00 GMT",
    "min_value": 172921.0387
}
```

Limitations & Implications

- There were several limitations worth considering such as limited coverage, data accuracy, sampling bias, and limited data variables.
- Our source, Zillow, may not cover all regions of the US comprehensively.
  - The data could be more robust in more populated areas, and less so in more rural areas.
- The properties listed on Zillow does not represent the entire housing market.
- There could be bias in the data due to certain properties or sellers that may be over or under represented on Zillow.
- Factors such as neighborhood characteristics, school quality, crime rate, and economic indicators can all affect housing prices but Zillow does not account for such factors.
- With this data, we can see study how housing prices could be compared to other societal issues, see how the 1% rule affects the profitability of rentals, and compare housing market data to homelessness rates in higher cost living areas.

Conclusions

- Based on the data we analyzed, both rental and housing costs have experienced a significant surge since February 2015.
- The data suggested that California is by far the most expensive state to rent, which is shown by the top 5 most expensive cities to rent all being located in California.
- The top 5 months with the highest increases in rental rates came in the year of 2021, we believe to be due to the pandemic.
- More than 43% of home values in the dataset had value increases in 2022, roughly 2 years after the pandemic first started.
- To sum our analysis up, both housing and rental prices have steadily increased, but both have seen a dramatic surge since the pandemic began.