PROPERTY APPRECIATION ESTIMATION AND RECOMMENDATION FOR STRATEGIC REAL ESTATE INVESTMENTS

Group members -

Abhiram Vasudeva, Armaan Ashfaque, Niyati Malik, Purva Tandel, Sushanth Thota, Vamsi Dath

UNDERSTANDING THE PROBLEM

The goal is to forecast property trends across cities, enabling users to make informed investment decisions based on past appreciation in specific zip codes. This is crucial in today's property market, where individual investors and corporations can benefit from data driven insights.

DATA COLLECTION AND PRE-PROCESSING

Datasets used -

- **Zillow Housing data**: Downloadable CSV Region-specific, shows past/present month-wise actual price values

  Size: 26338 x 305 (Rows x Columns)

RegionID	SizeRank	RegionName	RegionType	StateName	State	City	Metro	CountyName	2000-01-31	...	2023-11-30	2023-12-31	2024-01-31
84603	7958	60601	zip	IL	IL	Chicago	Chicago-Naperville-Elgin, IL-IN-WI	Cook County	232456.542741	...	343421.323487	344257.361587	344565.242611
84604	20351	60602	zip	IL	IL	Chicago	Chicago-Naperville-Elgin, IL-IN-WI	Cook County	271796.201836	...	292289.531820	289469.058508	285246.999747
84605	21722	60603	zip	IL	IL	Chicago	Chicago-Naperville-Elgin, IL-IN-WI	Cook County	NaN	...	332176.136008	329449.819693	324993.520639

- **Mortgage Rate dataset**,

  https://fred.stlouisfed.org/series/MORTGAGE30US

  Federal Reserve Economic Data (FRED) website.

  Size: 1297 x 2

DATA COLLECTION AND PRE-PROCESSING

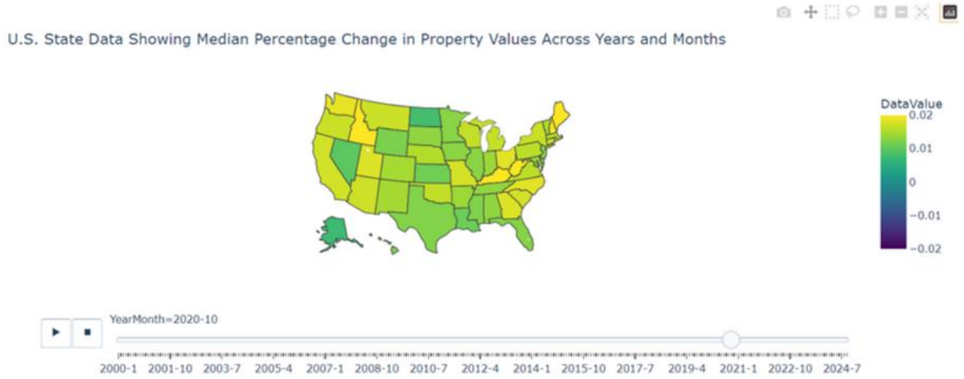
1. **Zillow CSV Files**
- Removal of Redundant Metadata and Columns and Handling Missing Values (NaN)
- Converting Datetime Columns to Numeric Values
- Calculating Monthly Percentage Change in Property Values
- Calculating Yearly Average Percentage Change

2. **Mortgage CSV File**
- Extracted Mortgage Rate Data
- Adjusting Mortgage Rate Data (Flipping Across the Z-Axis)

EXPLORATORY DATA ANALYSIS

We focused on answering several key questions that are crucial for real estate investors and stakeholders:

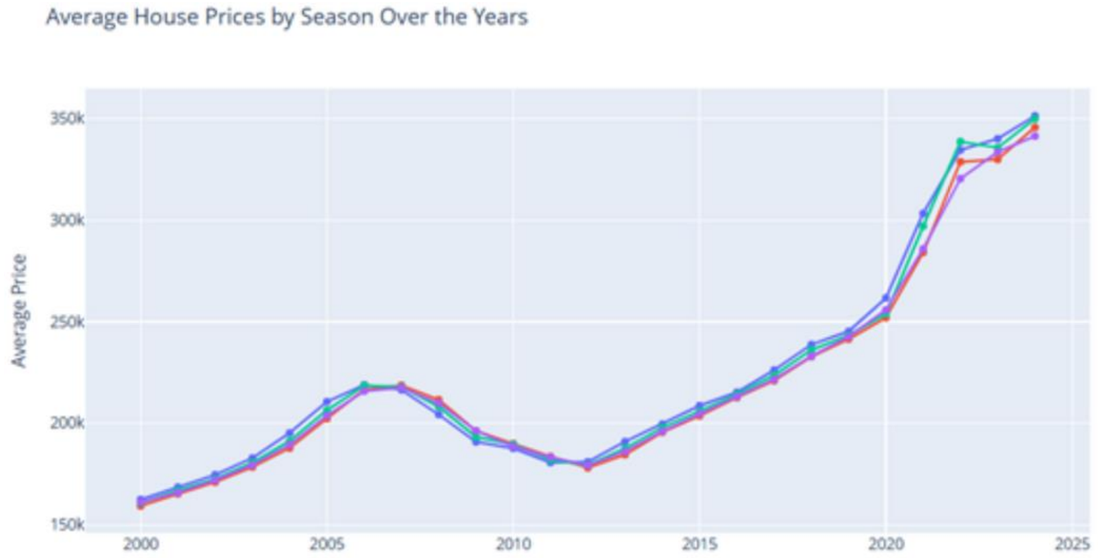
1. The influence of macroeconomic indicators, such as mortgage rates, on the average percentage change in property values across the United States.

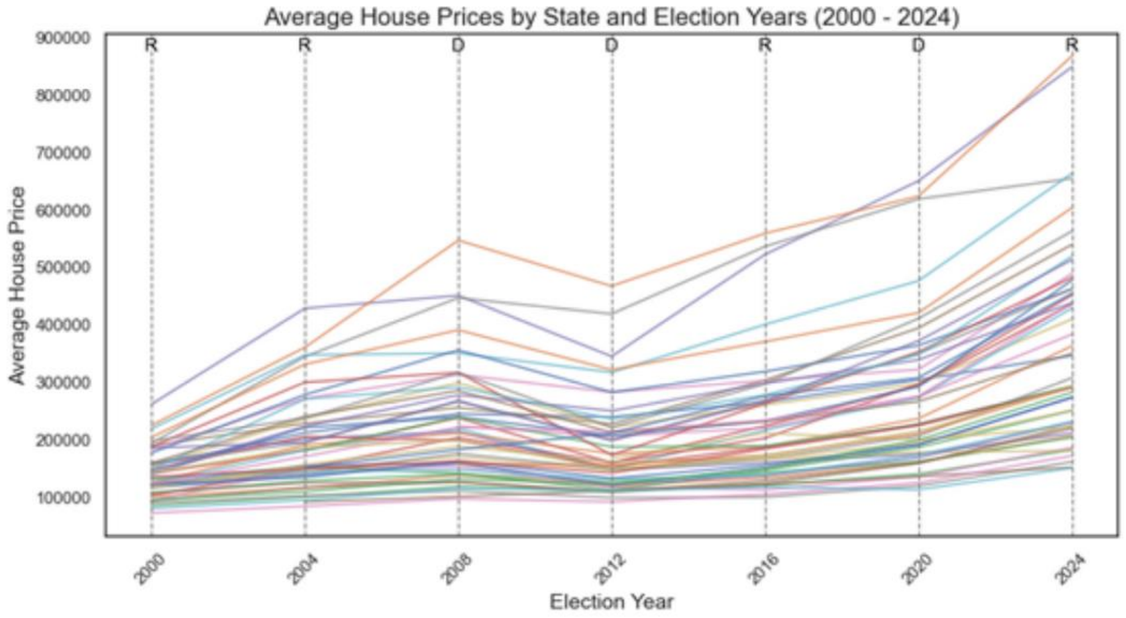


2. Seasonal trend analysis of changes in house prices over the past two decades, to identify recurring patterns or shifts in the housing market at different times of the year.

3. The impact of election results on housing prices, exploring whether political cycles and outcomes have a noticeable effect on the real estate market.

EXPLORATORY DATA ANALYSIS





DATA MODELING

Models used - Created a separate model for each zip code. -

- ARIMA
- Linear regression - by implementing lag features - Tried increasing the degree to 2 which led to minor decrease in error but increase in bias and variance, hence ignored.

Price	Date	Avg_Mortgage_y	lag_mortgage_rate	Price_t-1	Price_t-2	Price_t-3	Price_t-4	Price_t-5	Price_t-6	Price_t-7	Price_t-8	Price_t-9	Price_t-10
213807.536079	2001-01-31	7.0325	7.382	213093.513140	211931.255870	211069.140968	210694.883579	210924.920824	210892.224390	211651.051951	211804.731094	211769.525584	210785.302020
152364.398577	2001-01-31	7.0325	7.382	151024.307132	149501.590152	147739.390682	146044.282089	144593.453230	143431.480283	142330.265154	141138.728786	140122.689094	139146.069961
105085.828340	2001-01-31	7.0325	7.382	105008.778888	104617.189884	104219.472405	103927.949065	103570.880872	103334.339399	103147.149314	102956.005800	102909.028048	102918.703258
172369.059083	2001-01-31	7.0325	7.382	170178.368854	167654.850307	164868.438727	162409.717867	160452.105268	158612.701558	156797.662253	155344.147796	154007.712591	152144.240684
104272.941317	2001-01-31	7.0325	7.382	103982.496411	103447.591620	102963.876608	102604.911713	102245.015810	102056.679873	101957.960490	101821.442052	101846.094741	101928.968853
...	...	...	...	...	...	...	...	...	...	...	...	...	...
321416.544453	2024-09-30	6.1800	6.500	320764.218076	320634.651602	320524.648836	319889.190789	318021.383705	316342.572126	316036.662313	316960.510844	317575.939749	318408.474425
78644.070337	2024-09-30	6.1800	6.500	79027.439166	78758.257426	78545.548531	77494.125540	76810.852760	76292.584509	75943.183415	75736.976645	75298.402346	74260.344287
194649.933906	2024-09-30	6.1800	6.500	193784.325850	193407.728139	193009.236918	192174.483801	190328.300440	187934.494495	185437.183211	183994.040865	183644.769024	183797.381885
367017.449346	2024-09-30	6.1800	6.500	366492.471800	365094.615463	362540.939700	358966.135405	354618.405321	350853.871976	348974.319054	348768.561788	349379.606951	349013.713142
116354.445497	2024-09-30	6.1800	6.500	116111.785795	116207.287651	116293.470509	116251.151787	115614.674921	115125.502301	114732.321674	114850.862687	114699.743813	114610.009802

DATA MODELING

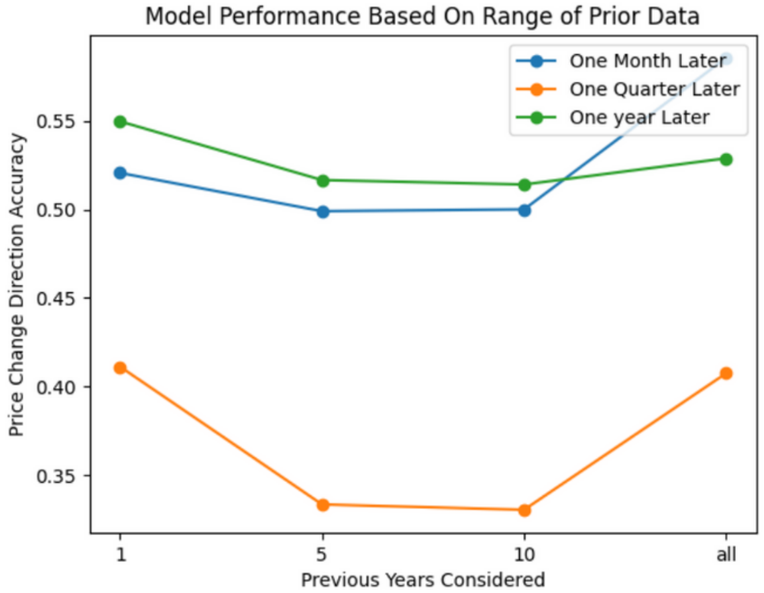
- Hypothesis testing - Does adding mortgage rate make any difference? Minor increase in evaluation metrics but not statistically significant!

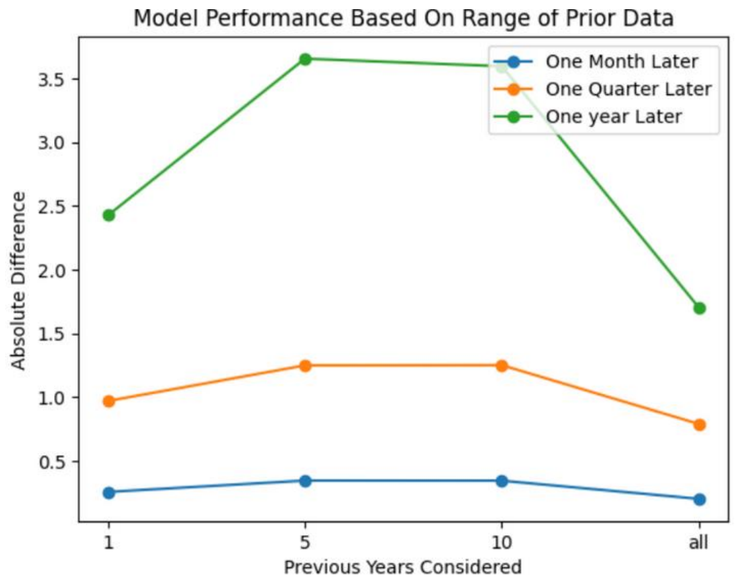
Metrics used for evaluation are the absolute difference in estimators and same sign accuracy.

RESULT

- One Month Sign Accuracy: 35%
- One Quarter sign Accuracy: 40%
- One Year Sign Accuracy: 47%
- One Month Absolute Difference Average: 55.48
- One Quarter Absolute Difference Average: 55.85
- One Year Absolute Difference Average: 55.59

RegionName	2024-10-31	2024-12-31	2025-09-30
77494	0.1	-0.5	-1.1
8701	0.5	0.1	-0.4
77449	-0.2	-1.1	-1.2
11368	-0.2	-0.9	-0.5
77084	-0.2	-0.9	-1.2





DECISION

Top 5 zip codes to invest in property to maximize ROI in 1 year are-

```
x =['RegionName', 'RegionType', 'StateName', 'State', 'City', 'Metro',
    'CountyName']
res[x]
```

✓ 0.0s

	RegionName	RegionType	StateName	State	City	Metro	CountyName
0	48505	zip	MI	MI	Flint	Flint, MI	Genesee County
1	61605	zip	IL	IL	Peoria	Peoria, IL	Peoria County
2	36610	zip	AL	AL	Prichard	Mobile, AL	Mobile County
3	71103	zip	LA	LA	Shreveport	Shreveport-Bossier City, LA	Caddo Parish
4	62914	zip	IL	IL	Cairo	Cape Girardeau, MO-IL	Alexander County

THANK YOU!

Any Questions?