# Let's Get Real-Estate!

GitHub: https://github.com/A-Sutaria/CS418_Project_FA23

| Arpan Sutaria | Devesh Patel | Jay Patel | Kevinkumar Patel | Htin Linn Htoo Than |
|---|---|---|---|---|
| asutar5 | dpate320 | jpate281 | kpate413 | hthan2 |
| Github: A-Sutaria | Github: dpatel9190 | Github: JayPatel47 | Github: kevu4567 | Github: HtinLinnHtooThan |

# Major Change

## Highway or My-Way

- Analyze chicago freeway traffic
- Insufficient data
- Unfeasible research

## Let's Get Real-Estate!

- Analyze real-estate market in the US
- Lots and lots of data!
- Interesting and practical research
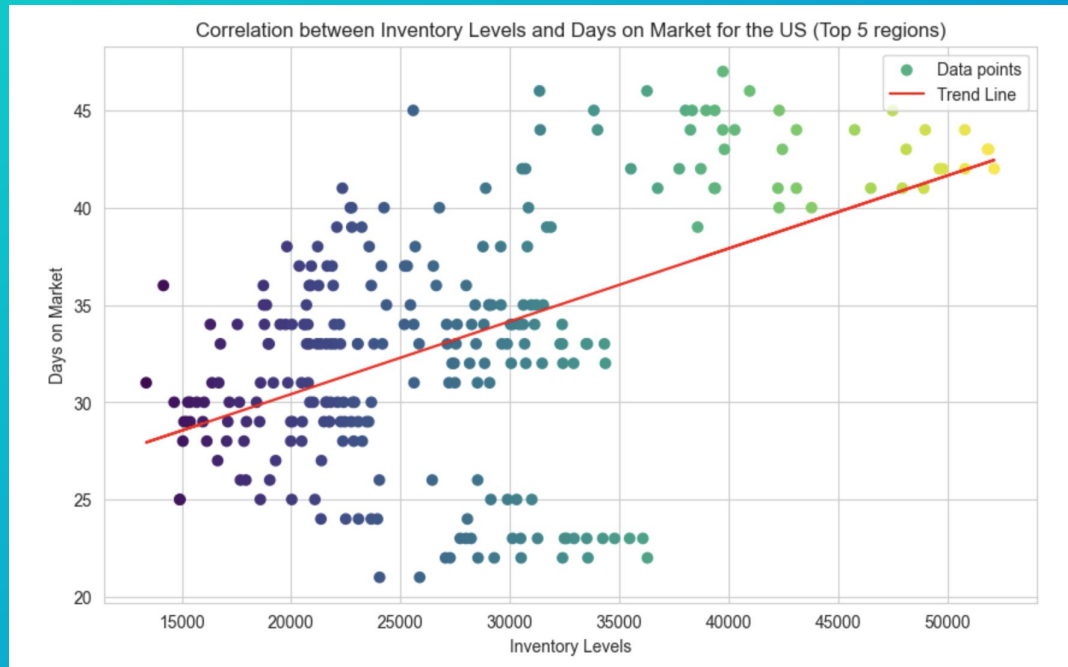
# Data

## Source
Zillow

## Type
Quantitative

## Range
January 2000 to September 2023

# Inventory level vs Days on Market

01



Correlation between Inventory Levels and Days on Market for the US (Top 5 regions)

- **Visualization shows a clear correlation between inventory level and number of days on market.**
- **Regions with higher inventory tends to have properties listed for a longer duration.**
- **This information is useful for stakeholders in real estate industry, including inventors, realtors, market analysts, as it provides insights into market dynamics.**

# Zillow Home Value Index vs Zillow Observed Rent Index

## 02



Scatter Plot of Average ZHVI vs Average ZORI for All Regions (Last 6 Months)
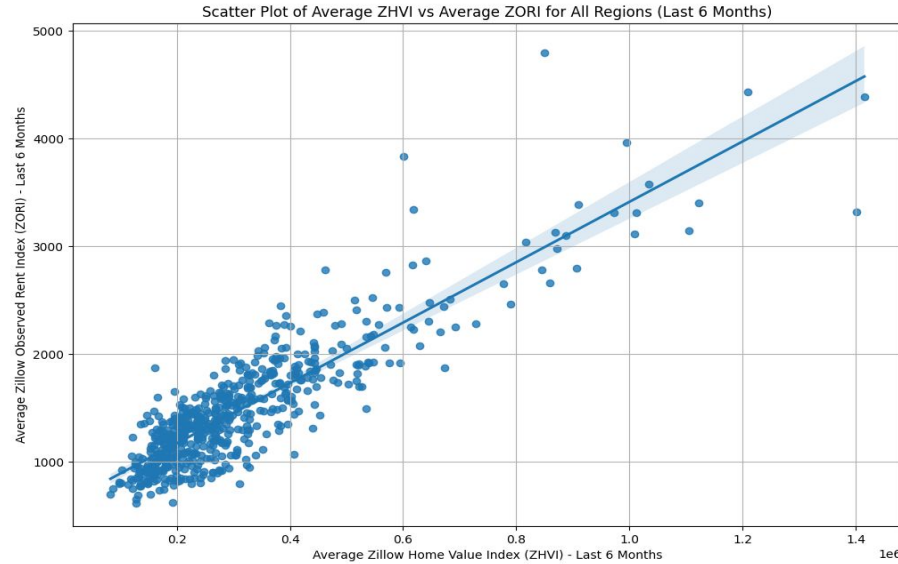
- Most regions show moderate rental pricing, while a few areas reveal higher rates, pointing to a distribution skewed toward luxury segments.
- **Who uses this data?** Real Estate Investors, Homebuyers and Renters, Economists and Analysts, and Real Estate Agents.

# Zillow Home Value Index (ZHVI)

The face of real-estate market

- **A measure of the typical home value**
- **Signifies market changes across a given region and housing type**
- **Reflects the typical value for homes in the 35th to 65th percentile range**
- **ZHVI can influence lots of other factors such as Zillow Observed Rent Index (ZORI), Days to Pending, Days to Close, etc.**
- **What if could predict ZHVI?**

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | RegionID | SizeRank | RegionNar | RegionTyp | StateName | 1/31/2000 | 2/29/2000 | 3/31/2000 | 4/30/2000 | 5/31/2000 | 6/30/2000 |
| 2 | 102001 | 0 | United Sta | country | | 121428.3 | 121642 | 121906.9 | 122475.1 | 123129.1 | 123830.3 |
| 3 | 394913 | 1 | New York, | msa | NY | 216219 | 217137.8 | 218065.1 | 219944.2 | 221890.1 | 224047.4 |
| 4 | 753899 | 2 | Los Angele | msa | CA | 222303 | 223130.3 | 224232.2 | 226424.6 | 228822.4 | 231203.3 |
| 5 | 394463 | 3 | Chicago, Il | msa | IL | 152289.7 | 152430.7 | 152699.2 | 153367.1 | 154170.6 | 155072.2 |
| 6 | 394514 | 4 | Dallas, TX | msa | TX | 125341.3 | 125397.2 | 125461.3 | 125628 | 125847.8 | 126070.2 |
| 7 | 394692 | 5 | Houston, T | msa | TX | 121045.2 | 121067.7 | 120983.2 | 121033.7 | 121080.3 | 121267.8 |

# Machine Learning Model

## 03

**Preprocess Data:**
- Imputed missing values using median

**Get User Input:**
- Prediction Horizon
- Test Dataset Size
- Whether to Scale Data

**Evaluate Model:**
- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- R2 Score

**Split Data:**
- Features X (historical data)
- Target y (future ZHVI values)

**Model Data:**
- GradientBoostingRegressor

# Conclusion

## 04

Predict changes in other features using ML predicted ZHVI

Assessing how external factors such as inflation, population, school districts, etc affect ZHVI, would lead to better predictions for ZHVI

Key details for property managers or even people looking to buy/rent properties.