### **Executive Summary**

This capstone project explores housing market trends and drivers across major U.S. metro areas from 2018 to 2023. The goal is to identify patterns in housing price growth and determine how external factors such as unemployment, natural disasters, income levels, and tax rates influence property values. The dataset is primarily based on Zillow’s housing metrics, enhanced with economic and environmental indicators.

### **Motivation**

The U.S. housing market has experienced major shifts over the past five years due to economic volatility, natural disasters, migration, and policy changes. Understanding the driving forces behind price fluctuations is valuable not only for real estate professionals and city planners but also for individuals looking to invest in or exit specific markets. With widespread access to data, this project aims to reveal which metros are most resilient, which are declining, and why.

### **Data Question**

1. How have home prices changed in major U.S. metro areas from 2018 to 2023?
2. Which metro areas experienced the fastest growth in home values?
3. What external factors (e.g., unemployment, natural disasters, tax rates) are most strongly associated with price changes?
4. Can we build reliable price forecasts for key metro areas for the next 1–2 years?

### **Minimum Viable Product (MVP)**

The final capstone deliverables will include:

* A cleaned and enriched dataset with Zillow housing data and additional public indicators
* Data exploration and correlation analysis between housing price trends and macro factors
* A Power BI dashboard displaying:  
  + Metro-level price trends
  + Growth rankings by metro
  + Interactive forecast visualizations
  + Filters by housing type and time
* A presentation highlighting the findings, intended for stakeholders such as real estate analysts or policy advisors.

### **Schedule**

| **Task** | **Target Completion Date** |
| --- | --- |
| 1. Get the Data | May 10, 2025 |
| 2. Clean & Explore the Data | May 25, 2025 |
| 3. Create Presentation / Dashboard | June 8, 2025 |
| 4. Internal Demos | June 25, 2025 |
| 5. Demo Day | July 10, 2025 |

### **Data Sources**

* [Zillow Research Data](https://www.zillow.com/research/data/) — ZHVI, Sales, Inventory (Metro level)
* [FRED](https://fredaccount.stlouisfed.org/) — Unemployment Rate, Median Income, Population (by API)
* [FEMA API](https://www.fema.gov/api/open/v2/DisasterDeclarationsSummaries) — Disaster Declarations
* [OpenICPSR](https://www.openicpsr.org/openicpsr/project/208462/version/V1/view) — Property, Sales & Income Tax Rates (2018-2022)
* Tax Foundation ([property](https://taxfoundation.org/data/all/state/property-taxes-by-state-county/), [sales](https://taxfoundation.org/data/all/state/2023-sales-taxes/), [income](https://taxfoundation.org/data/all/state/state-income-tax-rates-2023/)) — Property, Sales & Income Tax Rates (2023)

### **Known Issues and Challenges**

* Merging external data (unemployment, disasters, taxes) by metro area may require custom logic or manual alignment (e.g., mapping region names).
* Some data sources use state-level granularity instead of metro-level (e.g., tax data), requiring assumptions or state-to-metro mapping.
* Most of the data is taken from different sources, so you have to either combine or limit the Data Set.