

# Mike Hickey Capstone Proposal

## Executive Summary

This project aims to analyze real estate trends at the state level using a dataset from Redfin that includes key real estate data points such as median sale prices, number of homes sold, and inventory levels. To enhance the insights, this dataset will be joined with state population data, likely sourced from Census.org.

The motivation behind this project is to explore how housing market trends have evolved over time in relation to population changes. Understanding these trends can help identify patterns in housing affordability, supply and demand shifts, and regional market dynamics.

Key assumptions include the accuracy and completeness of the datasets, particularly in cases where historical data might be missing or inconsistent. One challenge will be aligning the timeframes and ensuring consistency between the Redfin real estate data and the population dataset.

## Motivation

My motivation for choosing this data project comes from a couple different factors. The first reason is that I am very familiar with real estate data, and I feel like I have a real contextual upper hand with being able to pull out meaningful insights. The other reason is that I really want to tailor my capstone to real estate so I can display it on my LinkedIn without raising any red flags with my current client / agent base.

## Data Question

The primary question driving this analysis is:

**"How have real estate trends across different states changed in relation to population growth or decline from the past 5 years?"**

Additional supporting questions include:

- Do states with rapid population growth see corresponding increases in home prices and sales volume?
- Are there states where population trends and housing trends diverge?
- How does housing inventory relate to population changes over time?

Relevant research includes studies on housing market trends, affordability crises in fast-growing regions, and reports from organizations like the National Association of Realtors and the U.S. Census Bureau. Similar analyses have been conducted at the city level, but this project aims to provide a broader state-level perspective.

Possibly bringing in State-to-State migration as well

## Minimum Viable Product (MVP)

### Minimum Viable Product (MVP)

The final capstone project will present a **state-by-state analysis** of real estate trends in relation to population changes from 2012-2024. The project will include:

- **Data Visualizations:**

# Mike Hickey Capstone Proposal

- Line charts showing trends in median sale prices and population over time.
- Map visualizations to show state price changes and population shifts.
- Scatter plots examining relationships between population growth rates and housing metrics.
- **Analysis Presentation:**
  - Power BI dashboard that houses the entire report as well as pages that show the SQL or Python code that was used so someone can see it completely in GitHub but also can see it within the project if they want to.
  - Some insight summaries.
  - Interactive elements.
- **Intended Audience:**
  - Real estate professionals, investors, policymakers, and researchers interested in housing market dynamics at a macro level.
  - Potential homebuyers or analysts looking for insights into affordability trends across states.

## Schedule (through <date of demo day>)

1. Get the Data (April 10, 2025)
2. Clean & Explore the Data (April 14, 2025)
3. Create Presentation of your Analysis (April 17, 2025)
  - It should be a presentation, but it could include a Jupyter Notebook or dashboard in Excel, Tableau, or Power BI
4. Internal demos (April 18, 2025)
5. Demo Day!! (April 25, 2025)

## Data Sources

<https://www.redfin.com/news/data-center/>

<https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html>

## Known Issues and Challenges

I need to find population data on a state level that has interesting data that relates to the real estate sales data I have.

From what I've seen on census.gov, I'll need to do a lot of cleaning to the population data to get it in the format I want and only the data I want.

See if state-to-state migration can be brought in effectively and looked at.