Tracking Sold Home Prices in the US

Jeremy Gormley

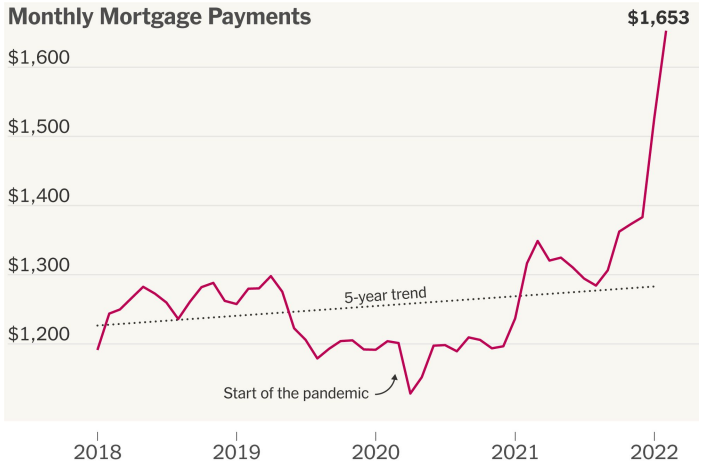
Outline

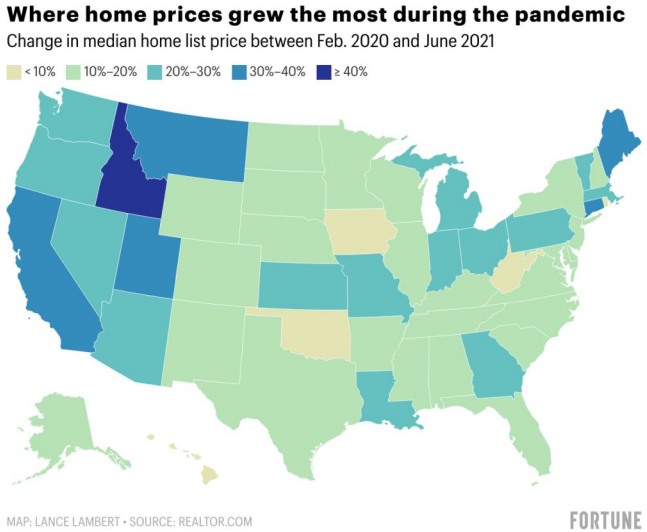
---

1. Introduction and purpose
2. Data and pipeline setup
3. App Demo
4. Next Steps

Introduction

- Home prices have reached record highs in the US since the start of the pandemic
- Regional disparities of home prices still largely exist
- Even within major metropolitan areas, home values often vary greatly





Purpose

- A (potentially) real-time tracker to determine in which direction home values are tracking
- Show where (by zip code) in each metro area prices tend to be greater/lower

Data Pipeline

INGESTION

-Weekly “Realty in US” API call using **Airflow**

-Approximately 200,000 data points

DATA STORAGE

-Store JSON data in MongoDBAtlas (BSON format) on Google Cloud

PROCESSING

-Clean and process data into usable format with Python Pandas. Data read into application from flat file.

DEPLOYMENT

- Create a web application with Streamlit. Deploy application to the web with Github/Streamlit Coud.

App Demo

Source code: https://github.com/jbg24/metis-homesale-project

Deployment: https://jbg24-metis-homesale-project-streamlit-app-2nhjq0.streamlitapp.com/

Next Steps

- Add more consistent updates via Airflow
- Increase number of analyses and introduce predictive modeling
- Include research-backed segregation metrics along housing sale prices