Assignment: Analyzing Influences on US Home Prices

Objective

The objective of this assignment is to find publicly available data on key factors influencing US home prices nationally and to build a data science model that explains how these factors have impacted home prices over the last 20 years. You will use the provided Python files to aid in your analysis and model building.

Key Factors to Investigate

You will focus on several key factors that influence home prices, including:

1. **Economic Factors**
    - Interest rates
    - Employment and income levels
    - Inflation
2. **Demographic Factors**
    - Population growth
    - Age distribution
    - Migration patterns
3. **Government Policies**
    - Tax policies
    - Zoning regulations
    - Subsidies and grants
4. **Housing Market Dynamics**
    - Supply and demand
    - Construction costs
    - Real estate market trends
5. **External Factors**
    - Natural disasters
    - Economic shocks

Steps to Complete the Assignment

Step 1: Data Collection

Find publicly available datasets that provide information on the key factors mentioned above. Potential sources include:

- Federal Reserve Economic Data (FRED)
- U.S. Census Bureau
- Bureau of Labor Statistics (BLS)
- National Association of Realtors (NAR)
- U.S. Department of Housing and Urban Development (HUD)

Step 2: Data Preprocessing

- Clean the data and handle any missing values.
- Normalize or standardize the data if necessary.
- Merge datasets to create a comprehensive dataset covering the last 20 years.

Step 3: Exploratory Data Analysis (EDA)

Use the provided Jupyter notebooks for guidance:

- **Main_Factors.ipynb:** This file contains a detailed breakdown of the main factors affecting home prices, including data and visualizations.
- **Demographic.ipynb:** This file examines demographic factors and their influence on housing demand and prices.

Step 4: Model Building

- Choose appropriate statistical or machine learning models to analyze the relationship between the key factors and home prices.
- Use the **Home_Value.ipynb** notebook for insights and methodologies on modeling home value trends.

Step 5: Model Evaluation

- Evaluate the performance of your models using relevant metrics (e.g., R-squared, RMSE).
- Validate the model using cross-validation techniques.

Step 6: Interpretation and Reporting

- Interpret the results of your models to explain how each factor impacts home prices.
- Use the **TCJA.ipynb** notebook as a reference for analyzing the impact of specific policies (e.g., Tax Cuts and Jobs Act) on home prices.

Step 7: Documentation

- Document your process, findings, and interpretations in a final report.
- Include visualizations to support your analysis.

Deliverables

- A comprehensive dataset that combines all relevant factors influencing home prices over the last 20 years.
- A Colab notebook containing your data preprocessing, EDA, model building, and evaluation steps.
- A final report summarizing your findings, including how each factor influences home prices and the performance of your models.

Submission

Submit the following:

1. Your Jupyter notebook(s) with all code and analyses.
2. The final report in PDF or Word format.
3. Any supplementary data files used in your analysis.