**Problem Statement:**

Find publicly available data for key \*supply-demand\* factors that influence US home prices

nationally. Then, build a data science model that explains how these factors impacted home prices over the last 20 years

**Data Collection**

I have collected data from publicly available sources to analyze the factors influencing US home prices. The data and their respective sources (website links) are listed below

* **Date** : These are the dates of observations (Quarterly observation from Jan-2000 to Apr-2023)
* **Unemployment Rate** : unemployment rate of USA from Jan-2000 to Apr-2023(Unit :Percent)

Source :<https://fred.stlouisfed.org/series/UNRATE>

* **GDP** :Gross Domestic Product value of USA from Jan-2000 to Apr-2023(Unit : Billion USD)

Source : <https://fred.stlouisfed.org/series/GDP>

* **Personal Tax** : Tax had to pay by house owner in USA during observed period(Unit :USD)

Source :<https://fred.stlouisfed.org/series/CXUSTATETAXLB1702M>

* **Fixed Mortgage rate** : Average of fixed mortgage rate in USA (Unit: percentage)

Source :<https://fred.stlouisfed.org/series/MORTGAGE30US>

* **Recession** : U.S. recessions as inferred by GDP-based recession indicator. This is Binary variable(1 -Recession ,0-No Recession)

Source :<https://fred.stlouisfed.org/series/JHDUSRGDPBRUNRATE>

* **Total Construction Spending:** : This variable gives value of spend in total construction in USA for the period we have taken for analysis. (Units : Million USD)

Source :<https://fred.stlouisfed.org/series/TTLCONS>

* **PERMIT** : Total housing permits passed in USA(Unit: Thousands of unit)   
   Source :<https://fred.stlouisfed.org/series/PERMIT>
* **Population** : Population of USA (Unit: Thousand) in observed period

Source :<https://fred.stlouisfed.org/series/POPTHM>

* **Life Expectancy** : This feature indicates the number of years a newborn infant would live

Source :<https://fred.stlouisfed.org/series/SPDYNLE00INUSA>

* **Housing CPI** : It is one of the most widely used indicators for inflation or price changes in an economy. This variable give value of Customer Price Index w.r.t housing for all urban consumers in USA

Source :<https://fred.stlouisfed.org/series/CPIHOSSL>

* **HPI** : The Home Price Index is a measure used to track changes in the prices of residential properties, specifically homes in USA This is our target variable in our dataset.

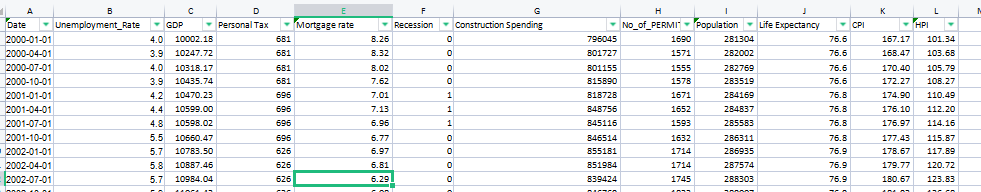
Source :<https://fred.stlouisfed.org/series/CSUSHPISA>

(The housing market is highly localized. Regional and local factors, such as zoning laws, local job markets, school quality, and neighborhood amenities, can often have an even more substantial impact on housing prices.But we have considered only national factors which are listed above for our analysis)

Other national factors like No.of marriage,immigrants also affect the housing price significantly. Due to unavailability , couldn’t include in our analysis.

Most variables in our dataset are reported quarterly. However, a few, such as Personal Tax and Life Expectancy, are annual. To align them with the quarterly frequency, I aggregated these variables by maintaining the same value across all four quarters within a year."

**Sample Dataset**



**Exploratory Data Analysis (EDA) has been conducted on the dataset mentioned above. Subsequently, a Linear Regression model was developed to predict HPI, and an ARIMA model was constructed for forecasting purposes.**