

Which countries are becoming more or less affordable for home buyers over time?

- Affordability Ratio = Avg House Price / Net Household Disposable Income
 - High Ratio = Less affordable
 - Lower Ratio = more affordable

```
#!/Users/tonydao/Documents/housingMarketProject/global_housing_market.csv
import pandas as pd
import matplotlib.pyplot as plt

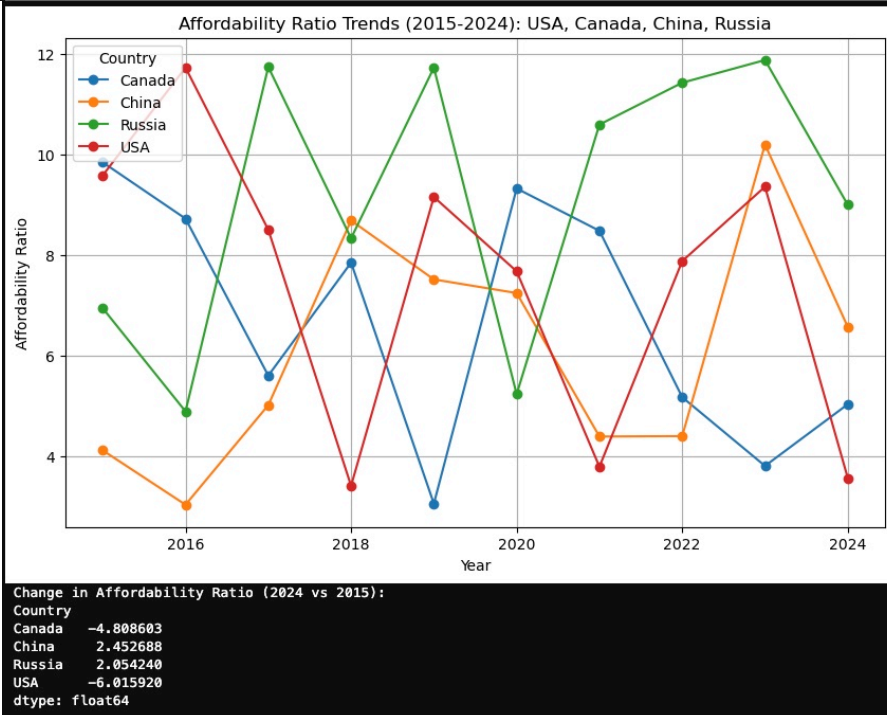
# Load the dataset
df = pd.read_csv('/Users/tonydao/Documents/housingMarketProject/global_housing_market.csv')

# Filter for the four countries of interest
countries = ['USA', 'Canada', 'China', 'Russia']
df_filtered = df[df['Country'].isin(countries)]

# Pivot the data for easier plotting (years as index, countries as columns)
pivot_df = df_filtered.pivot(index='Year', columns='Country', values='Affordability Ratio')

# Plot affordability ratio trends for the selected countries
pivot_df.plot(figsize=(10, 6), marker='o')
plt.title('Affordability Ratio Trends (2015-2024): USA, Canada, China, Russia')
plt.xlabel('Year')
plt.ylabel('Affordability Ratio')
plt.legend(title='Country')
plt.grid(True)
plt.show()

# Calculate the change in affordability ratio from 2015 to 2024 for each country
change = pivot_df.loc[2024] - pivot_df.loc[2015]
print("Change in Affordability Ratio (2024 vs 2015):")
print(change)
```



- Canada: Homes in Canada became more affordable between 2015 and 2024
 - USA: the USA saw a significant improvement in affordability over this period
 - Russia and China: Worsen affordability.
- How does the trends for GDP growth, inflation rate, and population growth look like for USA, China, Canada and Russia?

```

countries = ['USA', 'Canada', 'China', 'Russia']
df_sel = df[df['Country'].isin(countries)]

# --- 1. SCATTER PLOTS FOR EACH COUNTRY ---
fig, axes = plt.subplots(4, 3, figsize=(18, 20))
for i, country in enumerate(countries):
    cdata = df_sel[df_sel['Country'] == country]

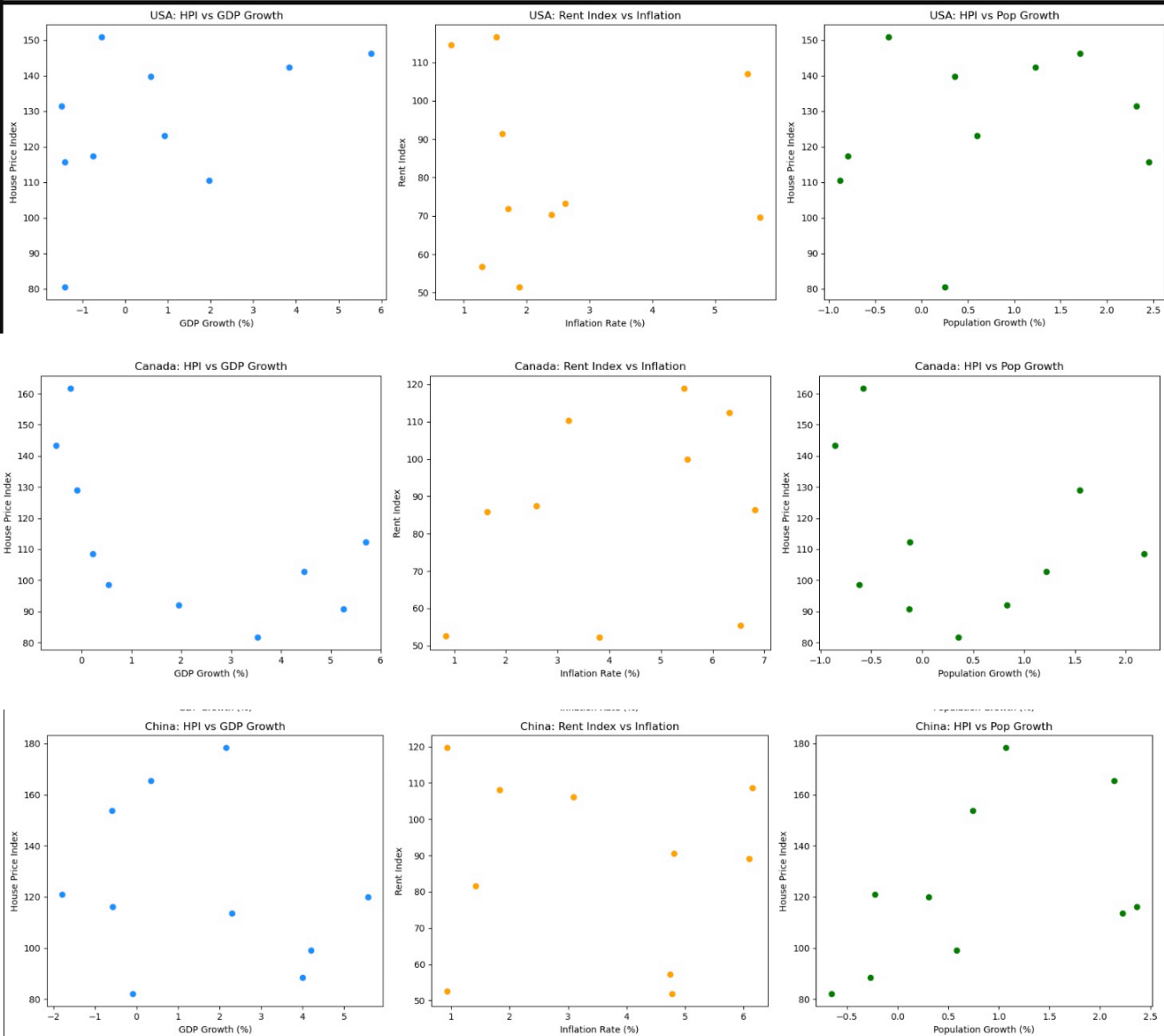
    # House Price Index vs GDP Growth
    axes[i, 0].scatter(cdata['GDP Growth (%)'], cdata['House Price Index'], color='dodgerblue')
    axes[i, 0].set_xlabel('GDP Growth (%)')
    axes[i, 0].set_ylabel('House Price Index')
    axes[i, 0].set_title(f'{country}: HPI vs GDP Growth')

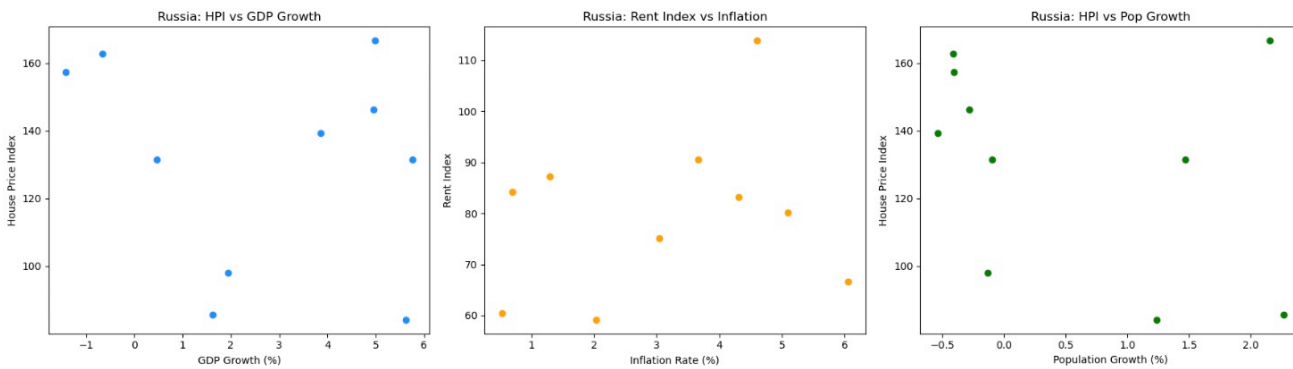
    # Rent Index vs Inflation Rate
    axes[i, 1].scatter(cdata['Inflation Rate (%)'], cdata['Rent Index'], color='orange')
    axes[i, 1].set_xlabel('Inflation Rate (%)')
    axes[i, 1].set_ylabel('Rent Index')
    axes[i, 1].set_title(f'{country}: Rent Index vs Inflation')

    # House Price Index vs Population Growth
    axes[i, 2].scatter(cdata['Population Growth (%)'], cdata['House Price Index'], color='green')
    axes[i, 2].set_xlabel('Population Growth (%)')
    axes[i, 2].set_ylabel('House Price Index')
    axes[i, 2].set_title(f'{country}: HPI vs Pop Growth')

plt.tight_layout()
plt.show()

```





○ USA:

- HPI vs GDP Growth
 - No clear upward or downward trend. No direct correlation with HPI over the years.
- Rent Index vs Inflation Rate
 - No strong linear relationship with inflation.

○ Canada

- HPI vs GDP Growth
 - Same as USA
- Rent Index vs Inflation Rate
 - Plots are wide spread.

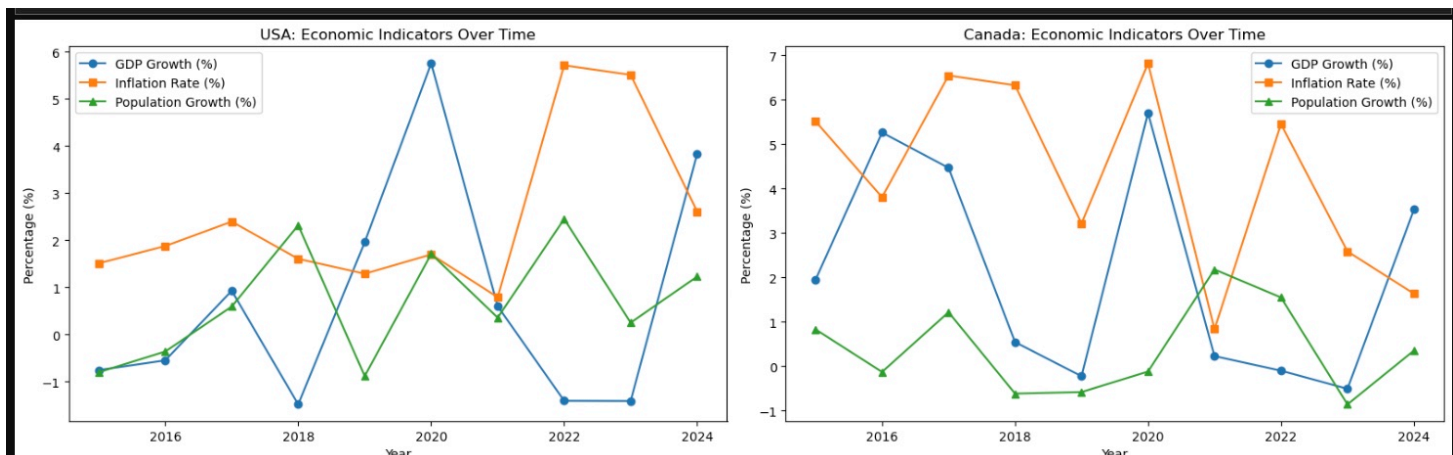
○ China

- HPI vs GDP Growth
 - No linear trend
- Rent Index vs Inflation Rate
 - Plots are wide spread

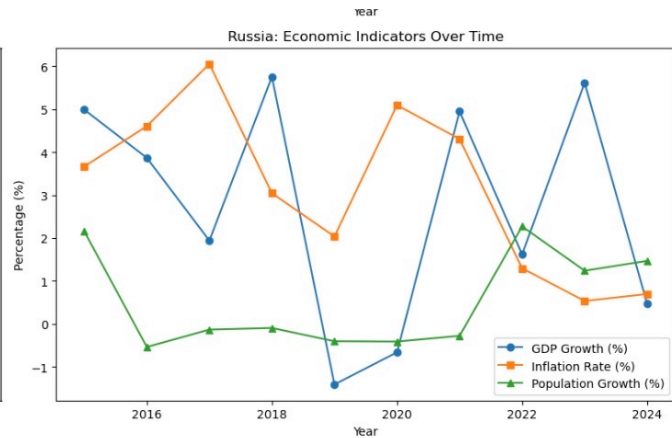
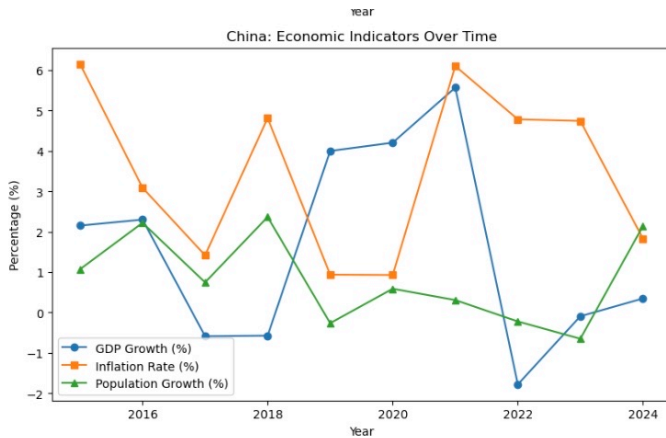
○ Russia

- House Price Index vs GDP Growth
 - No linear trend
- Rent Index vs Inflation Rate
 - Plots are wide spread

- Analyzing the GDP Growth, Inflation Rate and Population Growth For USA, Canada, China and Russia



- USA
 - "GDP Growth" decreased at 2020 but rebound later.
 - Inflation rate spiked in 2021-2022
 - Downward trend in population growth
- Canada
 - 2020 dip and recovery for GDP growth
 - good consistent positive growth



- China
 - slowly increasing GDP Growth
 - decrease in pop growth
- Russia
 - GDP Growth is volatile.
 - Pop Growth is stagnant.

• Correlation Coefficients Analysis

```
for country in countries:
    cdata = df_sel[df_sel['Country'] == country]
    print(f"\n{country} correlations:")
    print(cdata[['House Price Index', 'GDP Growth (%)',
                  'Inflation Rate (%)', 'Population Growth (%)', 'Rent Index']].corr())
```

```
USA correlations:
House Price Index    GDP Growth (%)    Inflation Rate (%) \
House Price Index    1.000000    0.470238    -0.615007
GDP Growth (%)       0.470238    1.000000    -0.369115
Inflation Rate (%)   -0.615007    -0.369115    1.000000
Population Growth (%) 0.203544    0.075369    0.359356
Rent Index           -0.333250    -0.338687    -0.012272

Population Growth (%)    Rent Index
House Price Index        0.203544    -0.333250
GDP Growth (%)           0.075369    -0.338687
Inflation Rate (%)        0.359356    -0.012272
Population Growth (%)     1.000000    -0.085252
Rent Index               -0.085252    1.000000
```

- House Price Index & GDP Growth: 0.47
 - Inc in GDP, HPI inc with it
- House Price Index & Inflation: -0.62
 - High inflation, Lower HPI
- House Price Index & Population Growth: 0.20
 - weak relationship
- Rent Index & GDP Growth: -0.34
 - moderate neg.
 - RI dec as GDP inc

Canada correlations:			
	House Price Index	GDP Growth (%)	Inflation Rate (%) \
House Price Index	1.000000	-0.601549	-0.113328
GDP Growth (%)	-0.601549	1.000000	0.351338
Inflation Rate (%)	-0.113328	0.351338	1.000000
Population Growth (%)	-0.278139	-0.020181	-0.113219
Rent Index	0.397677	-0.525043	0.294667
	Population Growth (%)	Rent Index	
House Price Index	-0.278139	0.397677	
GDP Growth (%)	-0.020181	-0.525043	
Inflation Rate (%)	-0.113219	0.294667	
Population Growth (%)	1.000000	-0.322858	
Rent Index	-0.322858	1.000000	

- House Price Index & GDP Growth: -0.60
 - Moderate neg
- House Price Index & Rent Index: 0.40
 - Moderate pos
- Rent Index & GDP Growth: -0.52
 - Moderate neg

China correlations:			
	House Price Index	GDP Growth (%)	Inflation Rate (%) \
House Price Index	1.000000	-0.225577	0.157927
GDP Growth (%)	-0.225577	1.000000	-0.086799
Inflation Rate (%)	0.157927	-0.086799	1.000000
Population Growth (%)	0.471033	-0.121200	-0.050519
Rent Index	0.370444	0.303921	-0.087705
	Population Growth (%)	Rent Index	
House Price Index	0.471033	0.370444	
GDP Growth (%)	-0.121200	0.303921	
Inflation Rate (%)	-0.050519	-0.087705	
Population Growth (%)	1.000000	0.472818	
Rent Index	0.472818	1.000000	

- House Price Index & Population Growth: 0.47
 - Moderate pos
- House Price Index & GDP Growth: -0.23
 - weak neg
- Rent Index & Population Growth: 0.47
 - Moderate pos
- Rent Index & GDP Growth: 0.30
 - Weak pos

Russia correlations:			
	House Price Index	GDP Growth (%)	Inflation Rate (%) \
House Price Index	1.000000	-0.203693	0.353338
GDP Growth (%)	-0.203693	1.000000	0.018135
Inflation Rate (%)	0.353338	0.018135	1.000000
Population Growth (%)	-0.343294	0.160096	-0.597253
Rent Index	0.265067	0.197108	0.253750
	Population Growth (%)	Rent Index	
House Price Index	-0.343294	0.265067	
GDP Growth (%)	0.160096	0.197108	
Inflation Rate (%)	-0.597253	0.253750	
Population Growth (%)	1.000000	0.073502	
Rent Index	0.073502	1.000000	

- House Price Index & Inflation: 0.35
 - weak pos
- House Price Index & Population Growth: -0.34
 - weak neg
- Rent Index & Inflation: 0.25
 - weak pos

Which Countries Show the Largest Gap Between House Price Growth and Income Gap?

```
import pandas as pd
import matplotlib.pyplot as plt

# Load my data
df = pd.read_csv('/Users/tonydao/Documents/housingMarketProject/global_housing_market.csv')

# Calculate the Income Index
df['Income Index'] = df['House Price Index'] / df['Affordability Ratio']

# Get the first and last year for each country
first_year = df.groupby('Country').first().reset_index()
last_year = df.groupby('Country').last().reset_index()

# Merge to align first and last year data
growth_df = pd.merge(
    first_year[['Country', 'House Price Index', 'Income Index']],
    last_year[['Country', 'House Price Index', 'Income Index']],
    on='Country',
    suffixes=('_start', '_end')
)

# Calculate growth rates
growth_df['House Price Growth (%)'] = ((growth_df['House Price Index_end'] / growth_df['House Price Index_start']) - 1) * 100
growth_df['Income Growth (%)'] = ((growth_df['Income Index_end'] / growth_df['Income Index_start']) - 1) * 100
growth_df['Gap (%)'] = growth_df['House Price Growth (%)'] - growth_df['Income Growth (%)']

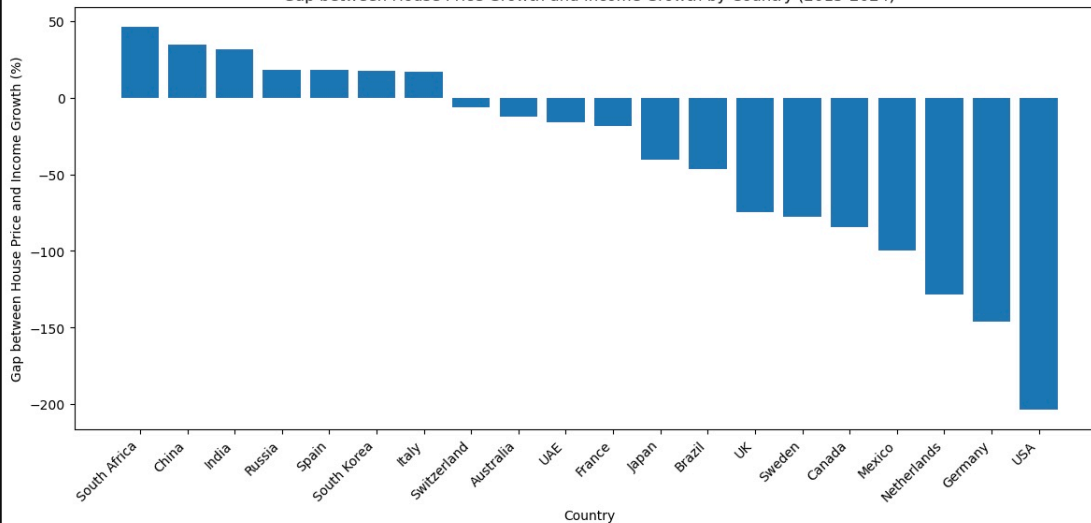
# Sort by the gap
growth_df_sorted = growth_df.sort_values('Gap (%)', ascending=False)

# Display top countries with largest gap
print(growth_df_sorted[['Country', 'House Price Growth (%)', 'Income Growth (%)', 'Gap (%)']].head(10))

# Plot
plt.figure(figsize=(12,6))
plt.bar(growth_df_sorted['Country'], growth_df_sorted['Gap (%)'])
plt.ylabel('Gap between House Price and Income Growth (%)')
plt.xlabel('Country')
plt.title('Gap between House Price Growth and Income Growth by Country (2015-2024)')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

	Country	House Price Growth (%)	Income Growth (%)	Gap (%)
12	South Africa	18.113260	-28.308555	46.421815
3	China	-7.235643	-41.841581	34.605938
6	India	-45.204115	-76.693735	31.489620
11	Russia	-21.081513	-39.091008	18.009494
14	Spain	-42.456103	-60.459684	18.003581
13	South Korea	-41.445911	-59.168468	17.722557
7	Italy	14.112240	-2.734319	16.846559
16	Switzerland	-2.585171	3.931192	-6.516363
0	Australia	-7.778166	4.577478	-12.355644
17	UAE	61.490928	77.509003	-16.018075

Gap between House Price Growth and Income Growth by Country (2015-2024)



- I calculated my income index by taking my HPI/Affordability ratio
 - USA has the largest negative gap, which is quite interesting because that would mean affordability has improved.
 - South Africa has the largest post gap meaning house price has grown faster than income.