**Uncovering Market Trends: Real Estate and Income Correlation in the Tri-State Area**

**Project 3 Team:** Zahra El Oula Kirati, Kyle McDaniel, Parth Pisolkar, Stephen Markovic & Genesis Ruiz

**Repository:** <https://github.com/parthpisolkar2002/real-estate-project3>

**Primary Objective:**

Examine the relationship between household incomes and real estate prices in select ZIP codes of the tri-state area from 2011 to 2021.

**Secondary Objective:**

Compare income levels and price per square foot between state capital cities and secondary cities to understand geographic influences on real estate trends.

**Data Collection:**

After selecting our datasets from Kaggle (USA Real Estate Dataset and US Household Income by ZIP Code), we explored and cleaned the data to include only the relevant information. We narrowed the USA real estate dataset to include only the tristate area and removed the column titled prev\_sold\_date. For the US Household Income data, we honed in exclusively on Household income, excluding categories such as Families, Married-Couple Families, and Nonfamily Households.

<https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset>

<https://www.kaggle.com/datasets/claygendron/us-household-income-by-zip-code-2021-2011?resource=download>

**Cleanup Process:**

1. Filtered data pertaining to the tri-state area.
2. Cleaned and combined datasets by utilizing ZIP codes as the merging key.
3. Leveraged the psycopg2 library within a Jupyter notebook to establish connections to PostgreSQL databases and transform the data into dataframes.
4. JSONified the data and accessed it using the Flask library.
5. Integrated user-activated visualizations into our webpage using the D3 JavaScript library.

**Data Analysis:**

After collecting and cleaning the data, we ended up with 33,000 unique zip codes in the tri-state area ready to be analyzed. We further narrowed our exploration to 5 zip codes near the capital of each state.

Albany, New York (Capital of New York) 12207, 12208, 12209, 12210, 12211

Hartford, Connecticut (Capital of Connecticut) 06105, 06106, 06107, 06112, 06120

Trenton, New Jersey (Capital of New Jersey) 08608, 08609, 08610, 08611, 08618

**Visualization:**

* **Line Graph: Displaying Income Levels in State Capitals**
* **Bar Chart: Comparing House Prices Across Different Cities**
* **Heat Map: Showing Real Estate Prices in the Tri-state Area**
* **Interactive Map: Illustrating Real Estate Listings Across the Tri-state Region**

**Insights & Conclusions:**

Based on the analysis conducted on the relationship between household incomes and real estate prices in the Tri-State Area, as well as the comparison between state capital cities and secondary cities, several key insights have been uncovered:

* **Geographic Advantage**: Proximity to major bodies of water and metropolitan areas significantly influences household incomes and real estate prices, indicating that geographic location plays a critical role in financial prosperity and property value appreciation.
* **Influence of Metropolitan Areas**: Cities near major urban centers like New York City experience substantially higher real estate costs, reflecting the strong economic influence of these metropolitan hubs on nearby property markets.
* **State Capitals vs. Secondary Cities**: The data shows that state capital cities do not always have the highest real estate prices compared to secondary cities, suggesting that factors other than political or administrative significance, such as local economic conditions and quality of life, impact real estate trends.
* **Regional Variations**: There are notable disparities in housing prices within the same state, with secondary cities like Stamford exhibiting significantly higher real estate values than their state capitals, underscoring the importance of regional economic dynamics and development patterns.

Overall, these insights highlight the complex interplay between geographic location, proximity to major urban centers, and regional economic conditions in shaping real estate markets in the Tri-State Area. Further research and analysis could delve deeper into these findings and explore additional factors influencing the real estate landscape in the region.