**Single Family Home Sale vs Rental prices in U.S.**

By Dataviz House Hunters

Team Members: Joby Augustine, Angelica Robles, Ra Ish Andrews

**Single Family Home sale and rental prices evaluation in U.S.**

Project analyses the Single-Family Residence Sale prices and the Rental prices in U.S. in a region, county or State.

Data Sets and Sources

Single Family Residence data (zillow house price data from kaggle.com)

<https://www.kaggle.com/paultimothymooney/zillow-house-price-data/version/14?select=State_Zhvi_SingleFamilyResidence.csv>

Single Family Residence Rental data (zillow house price data from kaggle.com)

<https://www.kaggle.com/paultimothymooney/zillow-house-price-data/version/14?select=City_Zri_SingleFamilyResidenceRental.csv>

**ETL Process**

ETL is performed using pandas with the steps below:

**Extract:**

* Datasets of Single-Family Residence and Single-Family Rental data for U.S. are used from [www.kaggle.com](http://www.kaggle.com)

Single Family Residence data (zillow house price data from kaggle.com)

<https://www.kaggle.com/paultimothymooney/zillow-house-price-data/version/14?select=State_Zhvi_SingleFamilyResidence.csv>

Single Family Residence Rental data (zillow house price data from kaggle.com)

<https://www.kaggle.com/paultimothymooney/zillow-house-price-data/version/14?select=City_Zri_SingleFamilyResidenceRental.csv>

* The CSV data sets are loaded into the data frames using Pandas read\_csv method.

**Transform:**

* The data for a year (June 2018 to May 2019) is pulled to the data frames for analysis. Data is filtered for one year and for each month.
* The data is cleaned, by eliminating the duplicates, and an aggregate column is added, finding the average sale price and average rental price to the respective data frames.

**Load:**

* The PostgreSQL database (SingleFamHouse\_SaleRentPrice\_db) and tables (sfhomesales and sfhomerental) are created with pgAdmin and are used to load the data into.
* Connection to the database is established and verified that the tables are created.
* The tables are loaded with the data in the data frames.
* The tables are queried in Jupyter Notebook to verify that the tables are loaded with the data.
* With PostgreSQL, the tables are queried in the pgAdmin to compare the Single-Family Residence sale price and the rental price for a specific Region, County or State.

**Analysis Conclusion:**

The analysis helps the Homeowners, investors and real estate agents in predicting the home sale/rental pricing, thus helping in planning the real estate business.

* Homeowners – help homeowners to plan their home selling /buying strategies.
* Rental providers/Investors – helps to adjust the rental amount as per the home price/rental price changes.
* Real Estate Agents – helps to get a pulse of the market.