*Texas Real Estate Listings Analysis and Takeaways*

Simple analysis on the listings of 500 ‘for sale’ houses, condos, townhomes, farms, and land in Texas.

References including the original data source and the link to the Tableau Public Dashboard for this analysis are provided at the end of this document.

*NOTE: information on data cleaning can be found in the README file, or in a relevant sql file available in the repository*

Setup:

* *Initial CSV file broken down into 3 tables with relational structure sharing one universal primary key of ‘id’. These tables were named ‘housing\_desc’, ‘house\_specs\_bath’, and ‘house\_specs\_size’*
* *Queries created for the following questions (Queries available in repository):*
  + *What is the average listing price for each combination of bathrooms and bedrooms present in the dataset?*
  + *How does price per square foot differ based on what year the property was built?*
  + *What is the average listing price, maximum listing price, and minimum listing price of properties within the range of 1890-2022, split into 27-year periods?*
* *Queries executed, then downloaded as CSV files for use in Tableau (due to Tableau Public restrictions on accessing databases)*
* *Fitting and readable dashboard created, no need for filtering options, as it is a very small set of data that I do not believe would benefit from adding filters.*

Takeaways from Tableau Visualization:

* *All but 3/65 0-bedroom, 0-bathroom properties were designated as farm/land. This is shown by the large count in the ‘Average Listing Price by Layout’ bar graph. Also visible is the single listing nearing $3m for the 4-bathroom, 6-bedroom property. Rather than being an outlier, this is more of a case of small sample size, and lack of other records with the same properties.*
* *Speaking in general terms, it seems the average listing price tends to increase as the ratio of bathrooms : bedrooms increases.*
* *Slight positive slope in linear regression model of ‘Price Per Square Foot Based on Year Built’, indicating the price per square foot is subtly increased for more recently built properties. This likely would have been more prevalent if not for outliers before 1900, and more accurate without our massive outlier in 1999.*
* *In the ‘Summary of Listing Prices Based on Year Built’ visualization, we can see the average listing price for properties built from 1890-1917 outweighs all other periods (Keeping in mind there are only 3 properties that fall under this category, one of which being an outlier). There is a dip in the average listing price in the 1946-1973 period. One could make assumptions about the quality of housing built during the WWII recovery period, but there is not enough evidence to make solid conclusions. We can see a generally widening gap between maximum and minimum listing price values as the ‘year\_built’ increases. I would attribute the minimum decrease to advancements in technology and materials usage. I would attribute the maximum increase to the buying power of individuals and their want for a more luxurious property (as these are technically individual cases when dealing with maximum and minimums).*

**References:**

**Link to original dataset:** [*Click Here*](https://www.kaggle.com/datasets/kanchana1990/texas-real-estate-trends-2024-500-listings)

**Link to Tableau Public Dashboard:** [*Click Here*](https://public.tableau.com/views/TexasRealEstateListings/TexasRealEstateListings?:language=en-US&publish=yes&:sid=&:display_count=n&:origin=viz_share_link)

**Link to Github Portfolio:** [*Click Here*](https://github.com/tre-rogers)