

San Antonio Real Estate Pricing Analysis

Justin Garza



Goal

As one of the fastest growing cities in the United States in a time when home ownership is at an all time low, San Antonio residents are in a very unique situation. For most people, purchasing a home is one of the biggest investments they will make in their lives.

Therefore, it is critical home buyers/sellers have as much information on the current market to ensure they are receiving the best deal based on their needs.

The goal of this project is to gain useful insight and perspective on San Antonio's current Real Estate Market in order to provide home buyers/sellers useful information to aid in their decision making.

Process

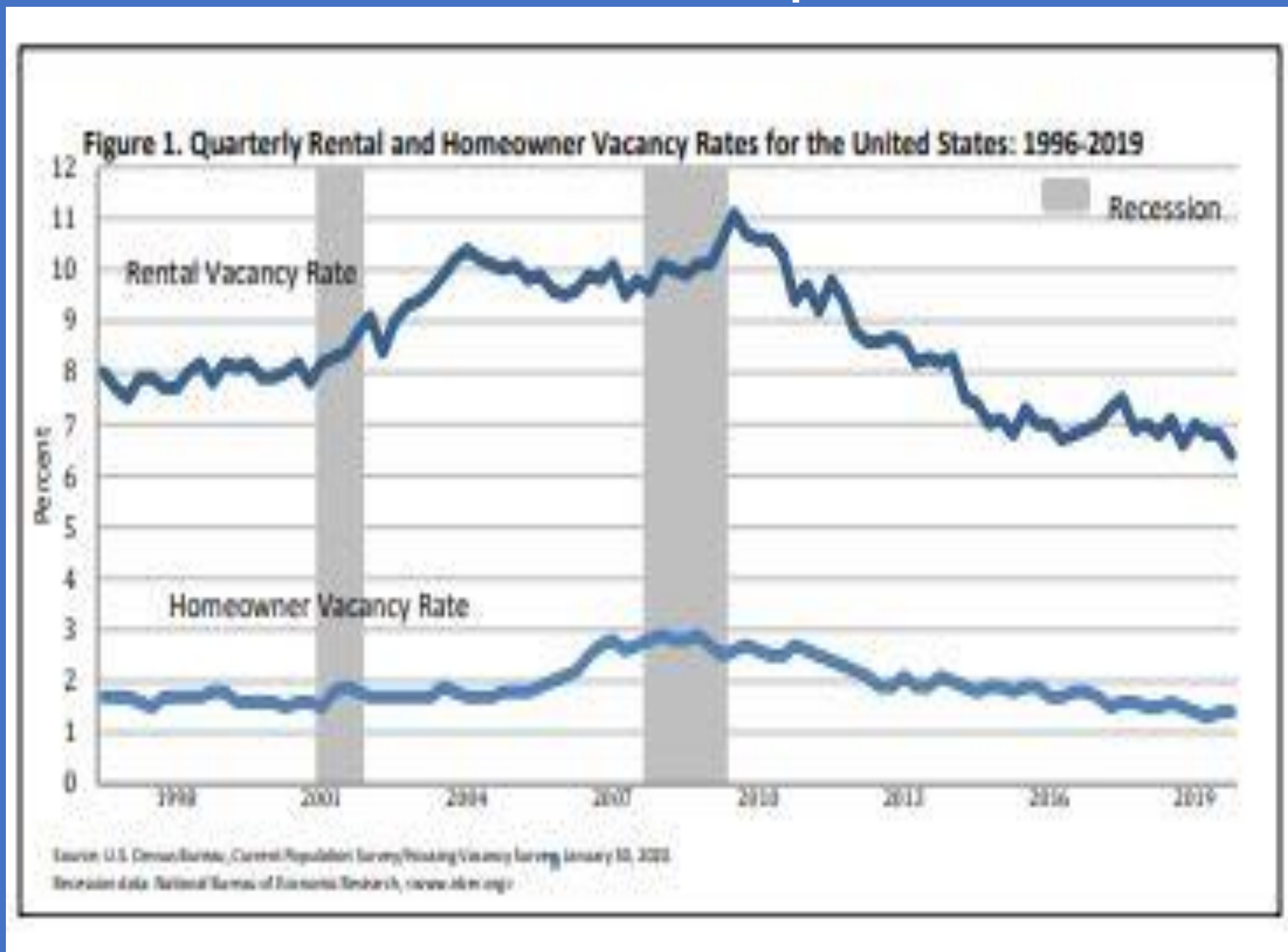
1. In order achieve the project goal data was web-scraped using R from Zillow.com by zip codes. Data captured included: Selling Price, Address, Square Footage, Number of Bedrooms, Number of Bathrooms.
2. This data was then Geocoded using Esri ArcGIS and a Hot Spot Analysis by Selling Price was performed to gain inference on the areas of Higher/Lower Priced Houses.
Additional data such as Urgent Care Centers, Law Enforcement Facilities, and Various Universities were also mapped to view relationships between these facilities and the Housing Hot Spot Regions.
3. Lastly, statistical modelling procedures were performed on the available housing attributes to predict Sales Price.

A Generalized Additive Model was chosen because of its easy-to-interpret qualities along with the high degree of accuracy it produces.

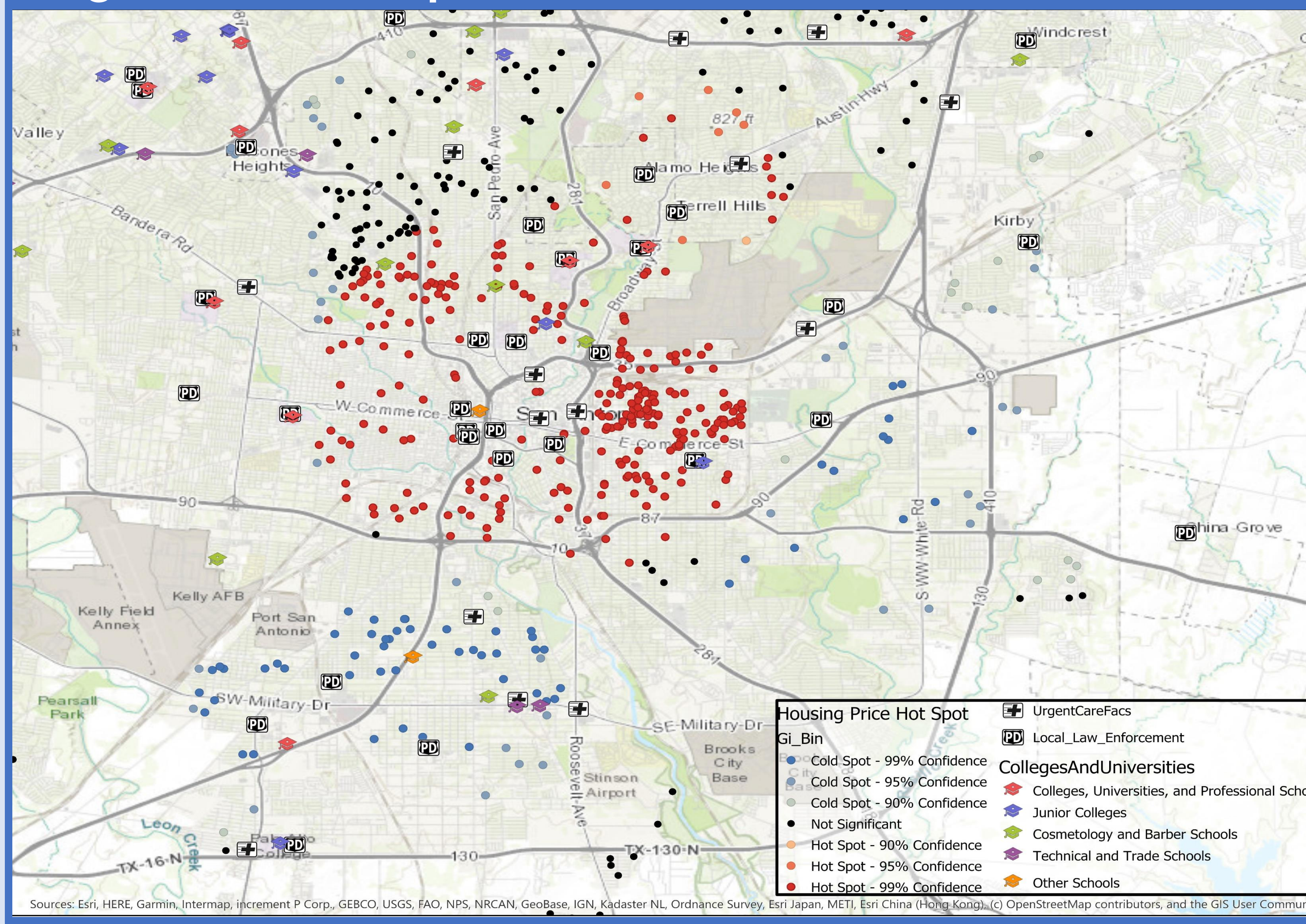
Results

- The Neighborhood Hot Spot Analysis shows us that higher Sales Prices are around the downtown area and pricing begins to diminish the further South & East the home is located.
- There appears to be a correlation between Higher Priced Homes and Law Enforcement Facilities, however this is most likely due to the area being downtown in high traffic areas.
- The Generalized Additive Model Proved to be statistically significant; explaining 71.8% of deviance within the model. Notably, we see that if bedrooms are added but bathrooms remain constant, there is a negative impact on Selling Price. However, adding bathroom and keeping bedrooms constant increases Sales Price.
- **Conclusion:** Although the data gathered proves to be beneficial in determining current Real Estate Pricing Trends within an area and an attribute-based level, more Real Estate Housing data (currently n=617) should be included to provide better results. Attributes such as Lot Size, Year Built, Story of House, Condition would greatly improve the statistical model and additional Housing Data Points would absolutely improve the Hot Spot Analysis.

United States Census Homeownership



Neighborhood Hot Spots

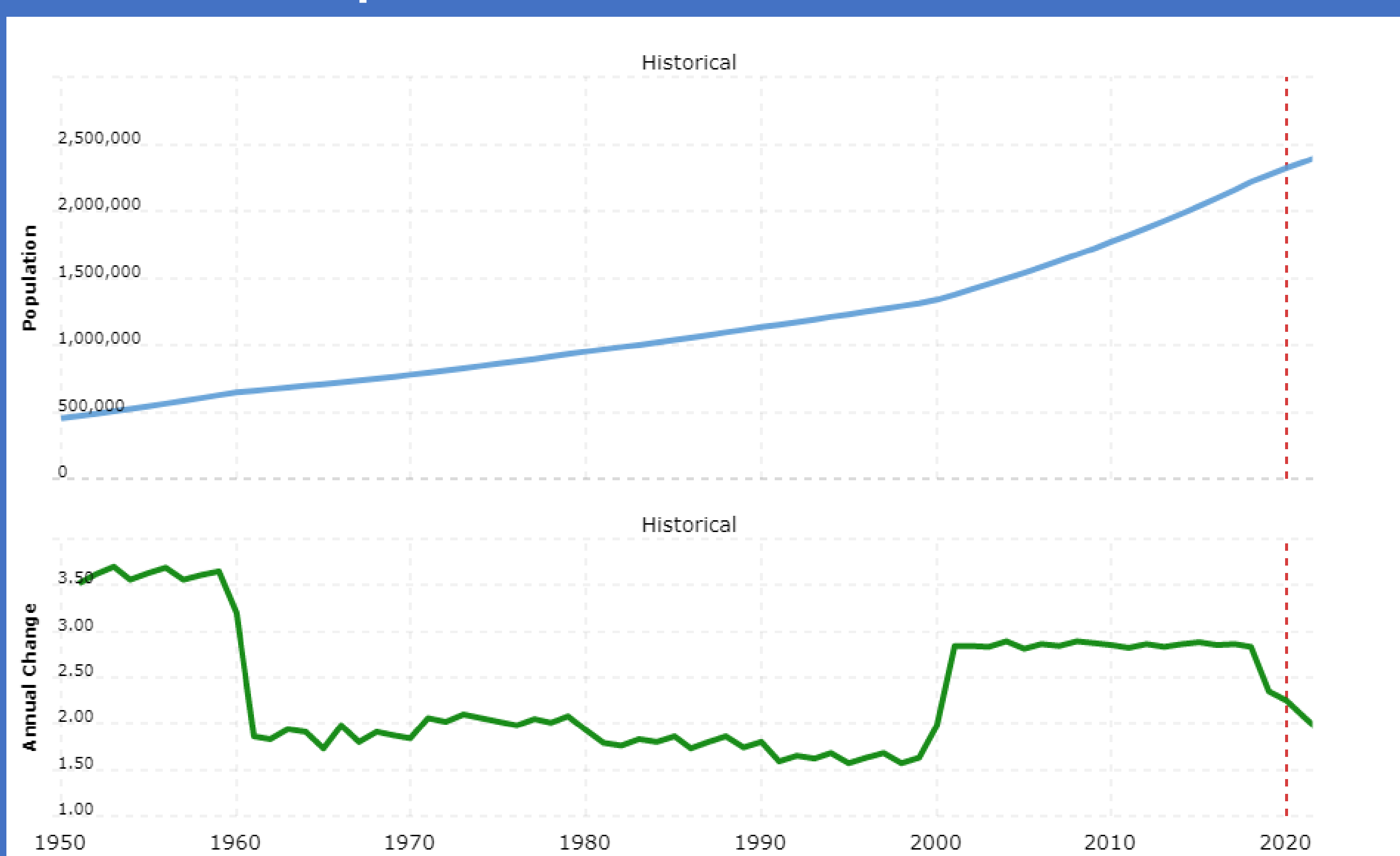


Model Summary

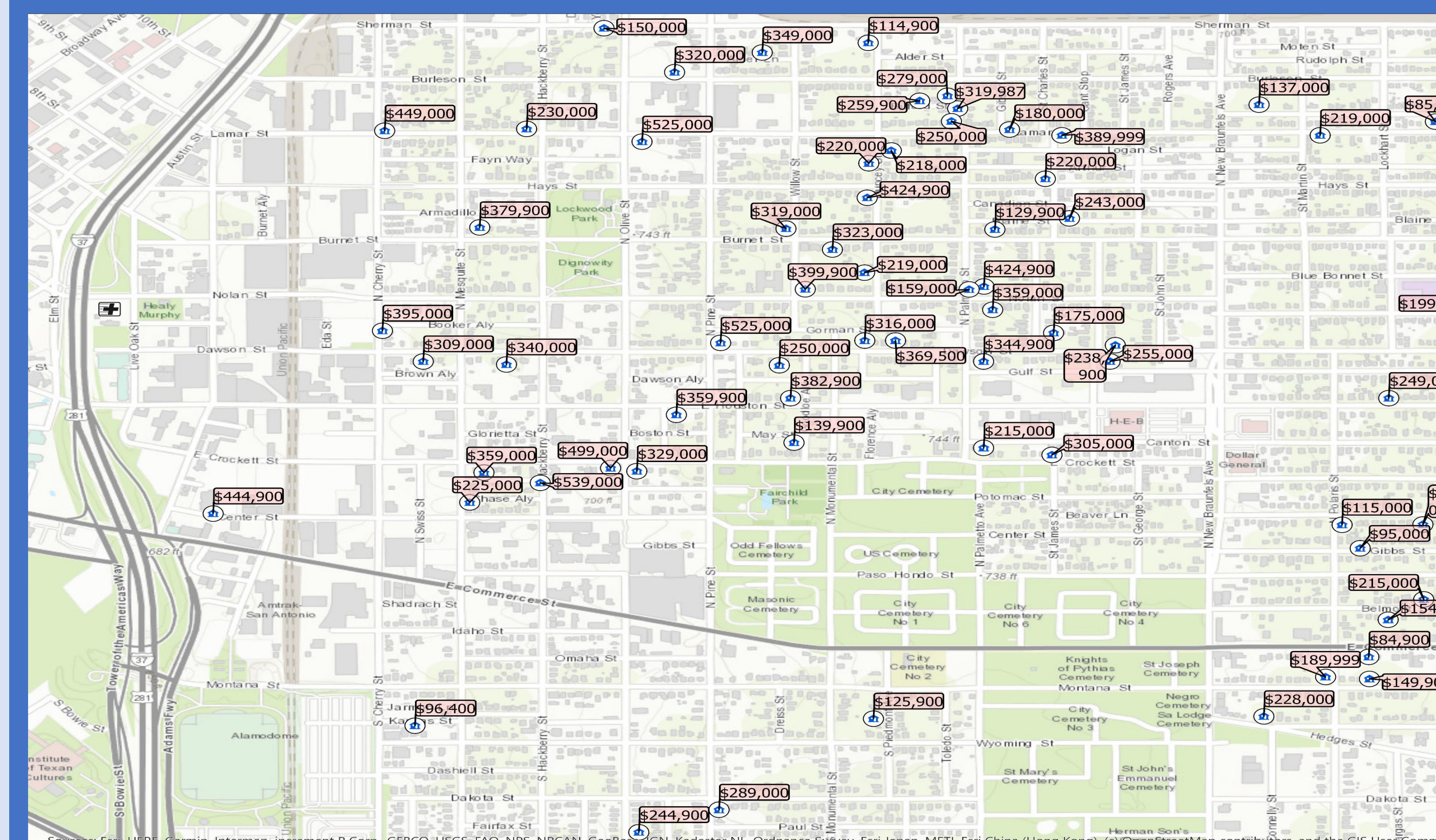
```
z.data<-read.csv("D:/zillow.data.final.5.csv")
gam.m2=mgcv::gam(Price~s(SqFt)+Bed+Bath,data=z.data, method = "REML")
summary.gam(gam.m2)

##
## Family: gaussian
## Link function: identity
##
## Formula:
## Price ~ s(SqFt) + Bed + Bath
##
## Parametric coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  424302      41440    10.239 < 2e-16 ***
## Bed         -104349      12394   -8.419 2.74e-16 ***
## Bath          98036      13587    7.216 1.61e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##              edf Ref.df    F p-value
## s(SqFt)      8.208  8.834 84.46 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj)  =  0.713   Deviance explained = 71.8%
## -REML       = 8457   Scale est. = 5.1725e+10   n = 617
```

San Antonio Population Growth



Sales Price



Model Output

